GC-TR-82-158.4(a)

⑫

A REVIEW OF
THREE-DIMENSIONAL VISION
FOR ROBOTICS

SPONSORED BY

DEFENSE ADVANCED RESEARCH PROJECTS AGENCY (DoD)
ARPA ORDER NO.:   3089
MONITORED BY: R. GOGOLEWSKI
UNDER CONTRACT NO.:   DNA 001-79-C-0208

DTIC
S   E
AUG 1 0 1982
H   D

PREPARED BY

GEO-CENTERS, INC.
320 NEEDHAM STREET
NEWTON UPPER FALLS, MA 02164

MAY 1982

82 08 09 116

GC-TR-82-158.4(a)

A REVIEW OF THREE-DIMENSIONAL VISION FOR ROBOTICS

ARPA ORDER NO.:   3089
PROGRAM CODE NO.:   9G10
PROGRAM ELEMENT CODE NO.:   62702E
CONTRACT NO.:   DNA 001-79-C-0208

The views and conclusions contained in this document are
those of the authors and should not be interpreted as neces-
sarily representing the official policies, either expressed or
implied, of the Defense Advanced Research Projects Agency
or the U.S. Government.

MAY 1982

GEO-CENTERS, INC.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER GC-TR-82-158.4(a)	2. GOVT ACCESSION NO. AD-A118055	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A REVIEW OF THREE-DIMENSIONAL VISION FOR ROBOTICS		5. TYPE OF REPORT & PERIOD COVERED (Interim, May, 1982) Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) GEO-CENTERS, INC. 320 Needham Street Newton Upper Falls, MA 02164		8. CONTRACT OR GRANT NUMBER(s) DNA 001-79-C-0208
9. PERFORMING ORGANIZATION NAME AND ADDRESS GEO-CENTERS, INC. 320 Needham Street Newton Upper Falls, MA 02164		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 3089; 9G10; 62702E
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency (DoD) 1400 Wilson Boulevard Arlington, VA 22203		12. REPORT DATE May, 1982
		13. NUMBER OF PAGES 56
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release - Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Robotics	Laser Scanning	Holography
Artificial Intelligence	Structured Light	
Optical Stereoscopy	Acousto-optics	
Proximity Sensing	Electro-optics	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The technologies appropriate to the collection and interpretation of three-dimensional data for robotics applications are reviewed. Included are optical stereoscopy, proximity sensing, laser scanning and structured light. Specifications for current systems are presented and areas of potential improvement are identified.

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73

EXECUTIVE SUMMARY

This report summarizes an examination of the technologies employed in acquiring three-dimensional information for robotic(s) applications. Of specific interest is the identification of technology concepts/ideas that have significant promise for improving abilities to acquire such information.

It became apparent during this study that acquiring three-dimensional information for robotics application can be usage-dependent. We have attempted to generalize this review and the conclusions reached. However, each prospective application should be carefully examined to identify the unique operating conditions or constraints which might be utilized to simplify the acquisition of three-dimensional information. In fact, at current performance levels of the state-of-the-art, the quantities of data associated with detailed three-dimensional information probably could not be effectively utilized. Imaging systems, although potentially capable of considerably enhancing robot performance, are expensive. The intelligent system designer should consider performing a trade-off between the dollars available and the acquisition of enough imaging capability to assure the efficient and timely completion of an objective. A proper trade-off allows consideration to be given whether expansion capabilities/capacities can be built into the system.

The technologies considered in this review include

- optical stereoscopy,
- proximity sensing,
- laser scanning, and
- structured light.

-ii-

GEO-CENTERS, INC.

In addition to surveying recent literature, facilities and researchers actively engaged in researching these technologies were contacted and queried. The information provided was examined with respect to known and anticipated requirements. Recommendations are made for both advanced research and extended efforts in the following general areas.

Optical Stereoscopy

Based on human vision and application of relatively simple triangulation theory, stereoscopy is receiving considerable attention for use in acquiring three-dimensional information. The most significant roadblock to effectively using stereoscopy in robotics is the problem of correlating two images to uniquely identify the same point in each image. The correlation problem is actively being addressed from several directions, including

- edge and vertex enhancement,
- grey-scale correlation, and
- shape correlation.

One avenue of approach not currently being addressed is the added use of color as a discriminant in aiding image correlation. Recent advances in acquiring color data with solid state image sensors add to the potential utility of the concept. Specifically, a trade-off study of grey-scale digitization levels versus color digitization levels should be undertaken.

Although improved solid state imaging capability is desirable for stereoscopy, the state-of-the-art in data collection is generally ahead of current abilities to effectively use the data generated. The commercial imaging industry (for TV, cameras, etc.) potentially represents a much greater driving force than robotics application for generating improvements in image sensor capabilities. However, some of the image correlation techniques presently being studied should potentially be implemented in hardware. Consideration should be given to coupling the image correlation techniques directly on the imaging sensor. Functional combinations, such as chemical sensors and microelectronics, are becoming more common and could be useful here.

GEO-CENTERS, INC.

Proximity Sensing

Although not three-dimensional, proximity sensing remains a considerable problem for robotics. Generally, as a manipulator, or perhaps even a moving robot itself, approaches an object under consideration, the usefulness of certain sensor systems greatly diminishes. This can be caused by obstructed views and/or inherent sensor limitations. The development of auxiliary proximity sensing techniques appears highly desirable. Recommended for more detailed study in proximity sensing applications are the use of fiber optic sensors and ultrasonic probing concepts and techniques.

Laser Scanning

The applications of laser techniques have been identified as one of the technologies exhibiting the greatest promise for versatile, three-dimensional data acquisition. Lasers can be used to acquire three-dimensional coordinate information in two distinctly different approaches. In one approach, ranging information is obtained from time-of-flight measurements; in the second, the unique character of the laser is used to generate a controlled illumination pattern which permits the acquisition of ranging information using simple triangulation.

Several ranging systems which have been formulated and used have shown great promise. For implementation in robotics, additional work is required in several areas. In the first application, data acquisition rates and signal-to-noise ratios could be improved with the development of higher power semiconductor lasers (preferably without cooling requirements). Semiconductor lasers are emphasized because of an inherent ruggedness and also because they are generally smaller and easier to handle. Second, improved means for nonmechanical laser beam deflection must be developed. Rotating/oscillating mirrors and/or prisms can perform the function, but they lack the ruggedness required for field or factory use. Acousto-optic deflection technology has use where beam deflection is extremely limited, but it cannot be used for the larger deflections desired for laser ranging (in principle, they could be serially used, but the resultant beam degradation and added computational complexity make such use impractical).

GEO-CENTERS, INC.

Structured Light

Utilizing a controlled laser illumination source with simple triangulation is a promising concept. The most advanced of these concepts generates an illuminating matrix of laser spots to be viewed with passive imaging technology. Presently, the most serious drawbacks to this technique appear to be array coding and the large quantity of data which must be stored and manipulated to generate the desired range information. There is an improvement in the amount of data over that required for the image correlation needed to perform stereoscopic analysis, but the numbers are still large. Improvements are needed.

With the recommendations given above, we would like to add one general observation. There are inherent strengths in the laser-scanning approach which should eventually be enhanced by advances in holographic techniques and in data storage, processing, and retrieval. It is not clear, at present, exactly what form such a hybrid system might take, but current research to develop real-time, erasable holographic storage elements should find application here.

GEO-CENTERS, INC.

TABLE OF CONTENTS

GEO-CENTERS, INC.

LIST OF ILLUSTRATIONS

GEO-CENTERS, INC.

1. INTRODUCTION

A robotic system can be described as one capable of receiving communications, understanding its environment, formulating and executing "plans," and monitoring its actions. Although both the capabilities and sales of "robots" show extremely sharp growth rates, only a limited number of systems are capable of performing all of the elements outlined above. Robots are finding increased utilization in applications too tedious, dangerous, and precise for human execution, and are proving to be more reliable, less demanding, and more cost-effective than human labor in many manufacturing applications. Increased robotics utilization is pushing their use into exploration and to other applications requiring decision-making capabilities.

Two of the robotic capabilities outlined above require the use of sensory systems to acquire data external to the robot. In both understanding its environment and monitoring its actions, a robotic system is dependent on sensors to probe and quantify the external environment. The sensors must accomplish these tasks accurately and rapidly.

Current sensor systems have limited abilities to acquire such information, and the primary technology employed (excepting tactile sensing) is two-dimensional imaging using conventional optical systems. Using some of the concepts described later in this report, the ability to acquire, process, and utilize two-dimensional information has been extended to permit the acquisition and use of limited amounts of three-dimensional information. However, current abilities to directly acquire three-dimensional information are minimal. This study was undertaken to review sensory systems and techniques for the purpose of identifying concepts and/or ideas having the potential to significantly enhance abilities to acquire three-dimensional range information.

GEO-CENTERS, INC.

The acquisition of three-dimensional information for application with robotic systems can be referred to as "range imaging." What is desired is the generation of accurate, three-dimensional coordinate maps which can be used

- for environment definition, and
- to quantify processes to be undertaken or which have just been completed.

Figure 1 summarizes the general techniques used to generate a coordinate map. Independent of the sensing technology employed, range information may be acquired either through a triangulation procedure, or by measuring the time for a signal to propagate from a source to a target and back ("time-of-flight"). Each of these techniques can again be subdivided.

```
                    ┌─────────────────┐
                    │  RANGE IMAGING  │
                    └─────────────────┘
            ┌──────────────┴──────────────┐
  ┌─────────────────┐           ┌─────────────────┐
  │  TRIANGULATION  │           │    TIME-OF-     │
  │   CALCULATION   │           │     FLIGHT      │
  └─────────────────┘           └─────────────────┘
     ┌──────┴──────┐               ┌──────┴──────┐
┌──────────┐ ┌──────────────┐ ┌──────────┐ ┌──────────┐
│STEREOSCOPY│ │  CONTROLLED  │ │ TIME-OF- │ │  PHASE   │
│          │ │ ILLUMINATION │ │ ARRIVAL  │ │  SHIFT   │
└──────────┘ └──────────────┘ └──────────┘ └──────────┘
```
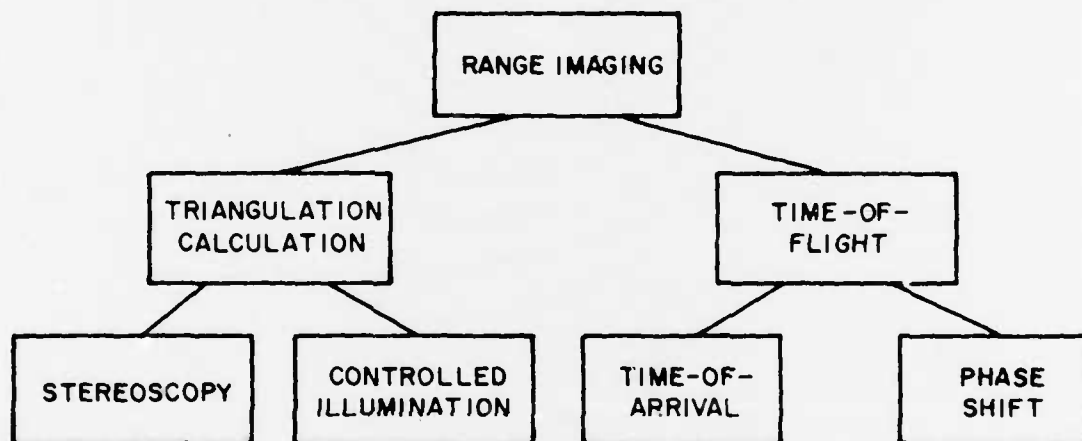
Figure 1. Range imaging concepts.

The triangulation technique can be divided into passive and active modes. In the passive mode, stereoscopy is accomplished using two separate imaging systems viewing an interrogation volume. Spatial coordinates are derived from a triangulation calculation which uses the coordinates of the "target" point in the image planes and the known parameters of the imaging systems. In the active mode, defined as controlled illumination, an imaging sensor(s) is (are)

GEO-CENTERS, INC.

used to view a volume which is illuminated by a controlled source. The illumination may provide a line, a point, or some other combination which either uses a symmetry of the problem or is based on a particular data processing scheme. In this approach, the known projection parameters of the illuminating source are used to constrain the problem and to reduce computational complexity.

Time-of-flight techniques generally employ colinear sources and detectors. As with the triangulation approach, this technique can also be divided into two modes. In the passive mode, a brute force approach, an impulse signal is generated and the propagation time is obtained as an elapsed time measurement. In the active mode, the source is modulated in a repetitive manner and the source and reflected signals are compared to measure a phase shift which can be interpreted in terms of range.

These techniques are reviewed in the remainder of this report and recommendations are made for additional work. Technologies specifically included are optical stereoscopy, structured light, ultrasonics, microwaves, and laser scanners. A limited number of miscellaneous concepts which could eventually be utilized are also included.

GEO-CENTERS, INC.

2. TRIANGULATION TECHNIQUES

To generate an accurate coordinate map with triangulation requires the determination of two direction vectors to a point. If these direction vectors are separated by a finite baseline, then a simple triangulation calculation can be used to determine the intersection and spatial coordinates of that intersection. Determination of the needed direction vectors may be accomplished using either passive sensor or using one active and one passive sensor.

The most common form of triangulation is passive stereoscopy, which uses two passive imaging sensors to acquire two-dimensional images of a volume of interest. The sensors are usually optical imaging cameras or arrays; direction vectors to a specific point can readily be generated by measuring the coordinates of that point's image in the image plane and then using the sensor optical parameters to calculate the direction vector.

In an alternate concept, one passive imaging sensor is replaced by an active sensor which can interrogate the volume of interest by a controlled illumination source. In this approach, while one direction vector is obtained from the passive sensor as before, the second is defined by the illumination source. By appropriately coding and controlling this source, the computational complexity of the problem can, in many cases, be reduced.

A spatial coordinate is generated from an estimate of the intersection points of a pair of direction vectors. The direction vectors are defined by the coordinates of a point in an image, or by the location of the illumination source, plus those system parameters which affect source direction or image signal direction.

GEO-CENTERS, INC.

2.1 STEREOSCOPY

The primary sensory system employed by animals, particularly humans, to acquire three-dimensional information is the eye. The human visual system employs two separated eyes to acquire stereoscopic imagery and exhibits excellent range-finding and object recognition capabilities. It is only natural then that we attempt to mimic this ability in robotic systems. The primary data acquisition technology being explored today for robotic systems is optical stereoscopy.

Our familiarity with the general concept, plus the fact that such a system would be passive (the only passive concept being pursued), make it quite appealing. With stereoscopy, the human eye is replaced by imaging optics and by an image sensor. The brain's reasoning and computational ability is replaced by hardware and/or software. Acquiring the desired three-dimensional information from two-dimensional images recorded by two, or more, optical systems is conceptually simple.

An outline of a stereoscopic sensor system is shown in Figure 2. With a knowledge of system parameters, a point can be located in space with a knowledge of its coordinates in each image plane. Estimating spatial location is a simple triangulation calculation which can be performed rapidly and accurately.
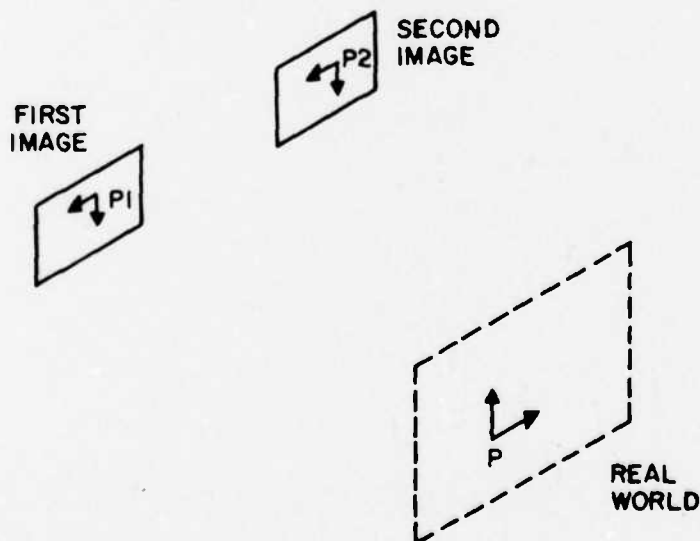
Figure 2. Stereoscopy imaging model.

GEO-CENTERS, INC.

The stereoscopic system may be divided into three separate elements: the optics, the image sensor, and the hardware and/or software required to convert the image data to three-dimensional information. Current abilities in optics will provide as much resolution as is needed for any known or anticipated stereoscopic system. The only detriment definable is cost. Commercial activities to develop less expensive and simplified photographic systems represent the key force in reducing this cost. Because of uncertainties, the direction vectors are more realistically conical and the spatial coordinate desired is somewhere within two intersecting cones. A least squares calculation is generally performed to obtain most probable coordinates.

As the two sensors are separated to increase the baseline, the conical error volume decreases until it minimizes at a relative angular separation of 45°. Beyond this point, the volume again increases.

All triangulation approaches have one serious drawback: with a bistatic system, both sensors may not necessarily be exposed to the same regions of a complex object. Obscuration and shadowing may make it impossible to develop coordinates for certain areas. Presently, there is no easy solution to this problem. Attention has focused on the image sensor and also on the hardware/software required to estimate spatial coordinates. Although not specifically germane to this study, the latter was included to permit a better understanding of the image sensor and its constraints.

The triangulation computation itself is not difficult or time consuming. The accuracy of the computation is dependent on the accuracy of the optical parameters and on the point image coordinates in the image plane. Optical parameters, which are usually known with a high degree of precision, will generally not be a limiting factor in achieving excellent triangulation results. A key limitation is the accuracy of the image coordinates used in the calculation; this accuracy is affected in two ways: 1) by the inherent resolution of the image sensor, and 2) by the accuracy with which a point can be uniquely identified in the two stereoscopic images. Ultimately, the latter constraint is the key element.

GEO-CENTERS, INC.

Manually, with photographically recorded images, coordinate information can be acquired with a high degree of precision. Photographic film has an inherently high resolution capability with an image readily divisible into millions of picture elements (pixels). Additionally, the eye and brain readily correlate two images to identify a common point with a high degree of accuracy. It is the latter which must be efficiently achieved in an automated fashion to permit ready usage in a robotic system. This difficulty has long been recognized and considerable effort is being devoted to developing both hardware and software approaches to successful image correlation. Grey-scale mapping, edge enhancement, vertex identification, and other techniques are being explored to accomplish image correlation. All of these techniques are dependent on the resolution ability and/or grey-scale capability of the image sensor. These capabilities are briefly reviewed here.

Both triangulation calculations and image correlation procedures require stable, well-registered sets of image data. Using camera systems with conventional photographic film as a recording medium readily satisfies this requirement, but the procedures required to generate useful data sets are both labor intensive and time consuming. For use in robotics, these operations must be automated and the image data sets must be obtainable as direct analog or digital electronic signals.

The simplest electronic image sensor which can be used for stereoscopy is the conventional television or video camera. Signal output, which is analog, must be converted to digital format for computer usage, but it is readily available and relatively inexpensive. Video cameras, recently extended in resolution ability (up to as high as 1000 x 1000 pixels), have two distinct drawbacks: 1) they tend to be relatively large in size, fragile, with significantly high voltage requirements; and 2) because of the electron beam sampling used to obtain data, physical image stability is not as high as desired.

For these reasons, the image sensor of choice for stereoscopy in robotics is the solid-state image sensor. Although there are several competing technologies for solid state image sensing, the most popular and advanced is the charge-coupled device (CCD), a silicon chip with a light-sensitive surface. A CCD, which can be manufactured in small size (postage stamp size is typical),

GEO-CENTERS, INC.

is a low-voltage device. It generates direct digital data and has fixed image registration. A typical CCD consists of a microscopic grid of light-sensitive elements etched or deposited on a silicon chip; each element converts light, striking it into an electrical charge. Pixel registration is, therefore, permanent and by careful mounting of a CCD pair, image-to-image registration can be well fixed.

Commercial imaging arrays are currently available at 256 x 256 elements (65,000 pixels) with 8-bit, grey-scale ability (256 levels). Arrays having double the number of elements along each axis (512 x 512) are now available; it is anticipated that 1024 x 1024 arrays will be available in the near future. The largest single problem with the current state-of-the-art appears to be picture element dropout and nonuniformity of pixel response across the array; both are being vigorously addressed.

The commercial video market provides the impetus for technology development. An indication of the commercial applications of this technology is the recent announcement of a magnetic video camera intended to replace the standard photographic camera (Appendix A). The magnetic video camera employs a 570 x 490 element array and an erasable magnetic video disc intended for playback and viewing with conventional color television sets. Such developments represent key changes in image technology. Only a fractional addition to this driving force is represented by robotics applications.

A sensor array may be duplicated in imaging ability by scanning a linear array across a field of view, or vice versa. In certain applications, such a technique may be well-suited (e.g., with component motion on a conveyer belt used to achieve scanning). Linear arrays are readily available now with densities as high as 2048 elements. Scanning must be accomplished with array motion, or with moving mirrors, and such designs lose generality. For this reason, two-dimensional staring arrays are preferred.

Although no recommendation is being made to support additional work in image sensors, or in optics in general, it is felt that two areas are worth consideration. Both areas are intended to address the image correlation problem and may ultimately impact image sensor concepts and fabrication techniques. In the first, it is felt that the use of color as a discriminant

-8-   GEO-CENTERS, INC.

should be considered in developing image correlation techniques. In the second, convolution and other computational approaches are being used for image correlation and it is felt that the relatively new technology of active acousto-optic processing, and other "on-chip" processing, may prove useful.

One of the techniques being explored to achieve image correlation is grey-level matching. This approach may prove particularly useful in industrial applications where the images considered have sharp discontinuous surfaces emphasized either by shading or by differences in angular reflectivity. Approaches being developed require significant grey-level discrimination and to date have proven difficult to implement. Since the human eye makes use of color as a discriminant, it is suggested that image correlation could be advanced by a similar use of color. Although effective correlation may require all eight bits of discrimination in monochrome images, perhaps with an added color discriminant the correlation can be accomplished at a lower quantization level. It is suggested that a trade-off study may provide interesting input to this hypothesis.

Should a trade-off prove the utility of using color, then imaging sensor technology would be directly impacted. Current technology produces a CCD with a monochrome response. Color response is generated with appropriate sequential filtration, either by filtering three separate CCDs, which leads to registration problems, or by sequentially moving filters in front of one CCD (also not desirable). Color separation must be accomplished on one CCD with adjoining or stacked pixels. However, it is felt here that the commercial sector probably will provide the primary impetus.

The suggested use of active electro-optic elements is prompted, first, by the realization that one approach to image correlation involves a convolution operation, and, second, by the observation that surface-acoustic waves can readily accomplish convolutions both rapidly and accurately (Appendix B). Although convolution operators are being developed with the understanding that they will be employed in a pipelined processing system, perhaps they can be more efficiently applied directly on the CCD chip. It is known that one application of this technology is permitting the efficient generation of Fourier transforms of an image both rapidly and accurately. This technology

-9-   GEO-CENTERS, INC.

should be explored in more depth, and should appropriate approaches be developed, then image sensor fabrication will be impacted. It is not clear that the commercial sector will provide a significant driving force in this technology, but a decision to proceed can await a successful demonstration of concept.

The most taxing application of passive stereoscopy is one which has images with very slowly varying grey-level content and no discernible edges or points which can be used for triangulation. With such conditions, passive stereoscopy may prove impossible or may not be feasible without producing significant errors.

2.2 CONTROLLED ILLUMINATION

Controlled illumination techniques involve the use of well-defined signal sources to scan a volume of interest. Because our ability to control optical signals is extensive, and because optical signals are minimally degraded over the generally short ranges required for robotics, the preferred technology for this application is optical. In general terms, a light source displaced from an imaging sensor is used in a controlled illumination mode. Both the form of the source and the manner in which it is used are controlled to maximize the data acquired and to minimize computational complexity. Typical light sources for this technology include light sheets, swept laser beams, laser spots, and other patterned formats. As opposed to the passive stereoscopy described previously, a range estimate is simplified because the dimensional and angular parameters (direction vector) of the source are well known.

A number of systems employing some or all of these techniques are currently being explored and developed. All have shown promise with respect to passive stereoscopy, but one particular system appears to have maximum potential (Appendix C). In this technique, a laser is used as the illumination source but its beam is modified in a unique manner. Double interferometry, using two shearing plates at 90°, is used to generate a rectangular array of controlled illumination beams. This array of beams, generated from one laser source, exhibits all of the positive characteristics of laser illumination in general and is readily controlled as a convergent, divergent, or parallel array.

-10-

GEO-CENTERS, INC.

The array is masked to control the number of elements (usually to a symmetrical array where the number of elements being used is a multiple of 2) and to space-code the array of spots, minimizing the amount of data needed to uniquely identify each one imaged.   Images of the space-coded array are sequentially obtained, and identification of a specific spot is accomplished by simple image subtraction.   As with passive stereoscopy, range estimates are then made by triangulation calculations.

GEO-CENTERS, INC.

3. TIME-OF-FLIGHT TECHNIQUES

Direct ranging can be accomplished by means of colinear sources and detectors to directly measure the time it takes a signal to propagate from source to target and back. Knowing the signal transport velocity, range is then readily calculated from the elapsed transport time. The most familiar use of this technique is standard sonar technology in which the echoes of acoustic pulses are recorded to provide reasonable range information.

As with the triangulation approach, the time-of-flight approach can be accomplished in two ways: 1) time of flight is directly obtained as an elapsed time when an impulse source is used, and 2) a CW source signal is modulated and the return signal is matched against the source to measure phase differences. These phase differences are then interpreted as range measurements.

Although optics (specifically lasers) is again the technology receiving the most attention, both ultrasonics and microwaves have application. The review performed here centers on the sensing signal used rather than the technology employed.

3.1 LASER SCANNING

A recent technological advancement which shows considerable promise for use in robotics is laser scanning. Lasers have been used extensively as range-finders, making use of single wavelength operation and minimization of beam divergence. Simple ranging is accomplished via time-of-flight measurement either between a laser source signal and a detector or with signals reflected from natural or man-made targets.

Although originally developed as a single-point measurement technique, the DoD has now pushed the technology into an imaging mode which can be used for range finding, target detection and identification, and moving target

-12-   GEO-CENTERS, INC.

indication. The laser radars (LIDARS) developed for this application are sophisticated units and have abilities which are being exploited in many new weapon systems. These systems, however, emphasize the longer range applications needed for fielded military systems. For the industrial sector, shorter range operation with even higher range and angular resolution capability is desired. The military laser scanner has, however, established feasibility and is providing a technological base to support the development of robotic sensors. Several such systems have been assembled and tested and development work to extend capabilities is ongoing (Appendix D).

Conceptually, the imaging laser scanner is well understood and with sufficient care, assembly can be successful. Generally, a laser source is used in a pulsed or CW mode to illuminate the desired target. In the pulsed mode, time-of-flight range gating is employed; in the CW mode, phase modulation with heterodyne detection is used for ranging. In the phase modulated CW mode of operation, an inherent range ambiguity results. Care must be taken to ensure that inferred ranges are not in error by the range quantum equivalent to the modulation frequency.

Although systems have been fabricated with range and angular resolution capabilities less than 1 mm (at 5- to 10-foot ranges), designers are now striving to achieve an order of magnitude improvement. Figure 3 identifies the major components of such a system. Each will be reviewed briefly with comments made on those which have potential for further development.

Operationally, the laser source must be considered (both type and wavelength) as well as the mode of operation (pulsed or CW), the beam scanning technology, and the detector type to be used. Of these, only the mode of operation is independent, although it is recognized that certain types of lasers lend themselves more readily to certain modes of operation. The CW heterodyne mode of operation is more difficult to implement but is potentially capable of greater range resolution. Angular resolution is essentially independent of operational mode and is dependent only on beam dispersion. The wavelength of the laser to be used must be carefully selected to ensure 1) maximum signal-to-noise ratios, and 2) simplicity of operation, ruggedness, and stability. It is recognized that specular reflections from edges and
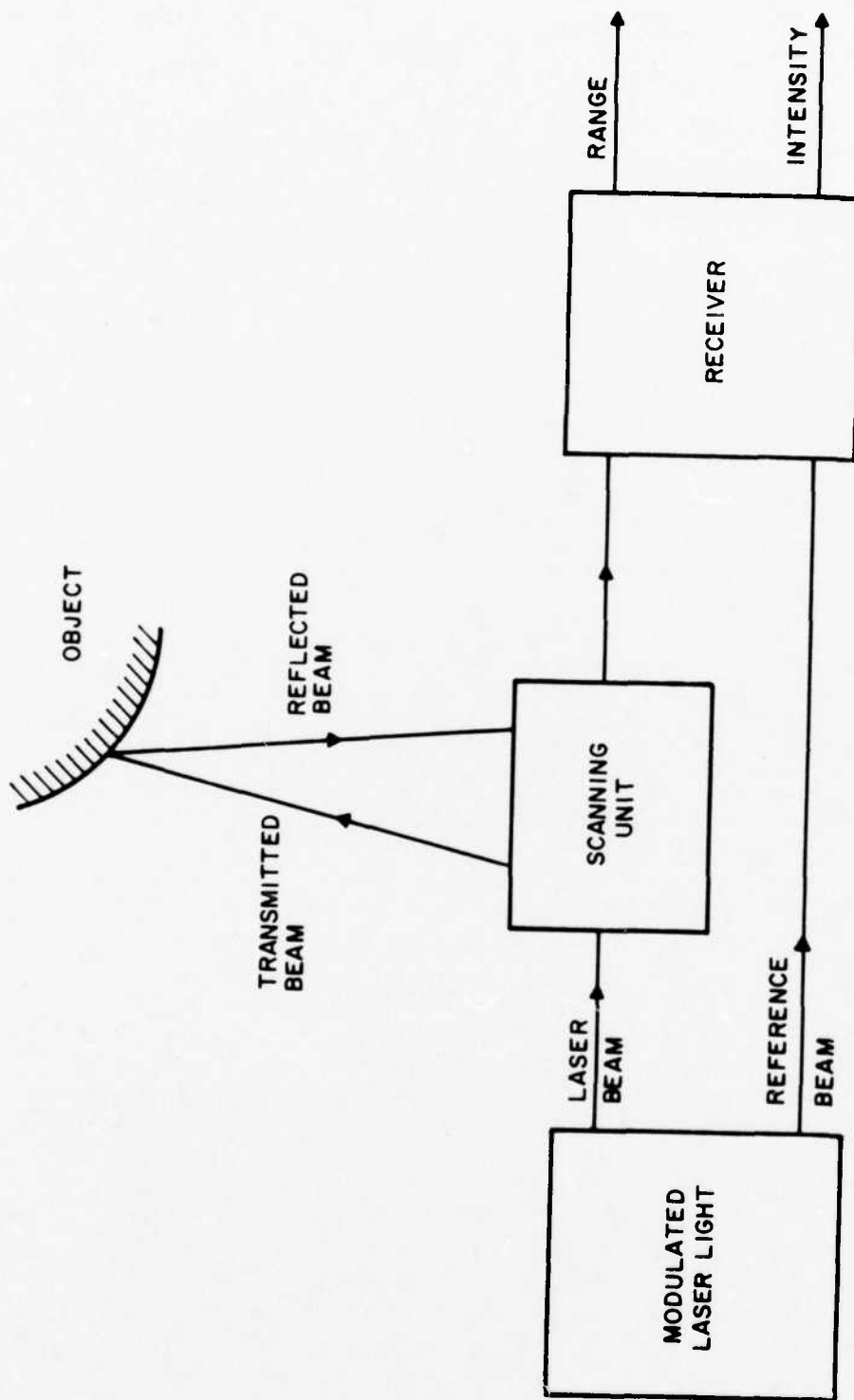
-13-   GEO-CENTERS, INC.

Figure 3. Laser Scanner System outline.

GEO-CENTERS, INC.

other discontinuities can deterimentally affect data acquisition. Such reflections are readily observed at the longer wavelengths whereas at the shorter, ultraviolet wavelengths all surfaces appear "rough" and specular reflections are less likely. This represents one definite advantage to using shorter wavelength sources.

There is a trend toward more compact, lower-cost laser systems. It is felt, for example, that 6-inch long helium-neon lasers will shortly be available. These small, stable, multimode lasers will have increased use in battery-powered portable scanning units. There are comparable advancements being made in semiconductor laser technology. These are of special interest for the application of laser scanning systems to robotics.

Standard techniques for beam scanning use moving mirrors and rotating prisms. Several new technologies have recently shown increased promise. These include holographic and acousto-optic techniques. Neither shows any current advantage over more standard techniques for robotics application.

The development of holographic scanners has been a significant recent advancement. This technology has been advanced by the commercial sector, primarily for data acquisition systems such as point-of-sale product code scanning. In this application, a spinning disc, containing a number of transmission holograms, is used to deflect and focus a laser beam by diffraction. Efficiencies of these holographic scanners have exceeded 90%, and show considerable promise for replacing rotating polygon spinners for the same purpose.

Rotating polygon spinners are also used for beam scanning but must be manufactured to extremely tight tolerances. Typical requirements are fractional wavelength flatness per surface and a surface-to-surface orientation tolerance of less than several arc seconds. New techniques in diamond point machining and on-line measurements are making these objectives attainable, but polygon elements are still extremely expensive. Holographic scanners, on the other hand, could be replicated very inexpensively by holographic recording techniques.

GEO-CENTERS, INC.

Acousto-optic beam scanning capitalizes on the fact that the index of refraction can be changed by applying pressure to the crystal. In Figure 4, the entering laser beam will be diffracted by the crystal. As pressure is applied to the crystal, its refractive index will change and the laser beam deflection will be modified accordingly. In practice, pressure is applied to the crystal through a piezo-electric material. By modulating the driving signal sent to the piezo-electric material, the laser beam is deflected. Although advances in this technology have been dramatic and useful, acousto-optic modulation for beam scanning is still limited to small fractions of a degree. The applications envisioned here would require tens of degrees of deflection, while still maintaining beam integrity.

POSSIBLE DIFFRACTED OUTPUTS

RF DRIVER

PIEZOELECTRIC TRANSDUCER

BONDING LAYER

INPUT LASER BEAM

THE PERIODIC GRATING IS ACOUSTICALLY INDUCED CAUSING REFRACTIVE INDEX CHANGES IN MEDIUM

SONIC ABSORBER

INTERACTION MEDIUM TYPICALLY A CRYSTAL OR DENSE GLASS

BRAGG ANGLE θ

- 2nd ORDER ETC.

-1st ORDER

oTH ORDER

+1st ORDER (DESIRED BEAM) (UNDIFFRACTED)

+2nd ORDER

+3rd ORDER ETC.

Figure 4. Outline of Acousto-Optic Laser Beam Diffraction.

Scanning with moving mirrors (galvanometers or resonant scanners) remains one of the easiest technologies to implement and is the least expensive. However, significant limitations are placed on the performance of such systems

-16-

GEO-CENTERS, INC.

by the inertial mass of the oscillating mirror. Current technology allows operation up to ~500 Hz, but advances in new lightweight substrate materials will allow operation at higher limits. The technology is still fragile, however, and not the most desirable.

3.2 ULTRASONICS

Active ultrasonic interrogation is regularly used to acquire accurate ranging information. The ultrasonic range finders used on some of the newer camera systems, for example, are capable of 0.1-foot resolution in the range of 5 to 35 feet. However, the beam width of the emitting source is almost a full 20°; therefore, the system angular resolution is limited.

The use of ultrasonics for imaging or range-finding has two inherent limitations. First, ultrasonic signals are severely attenuated in air with 1 dB/m being readily observed. In a fluid medium, attenuation is not as severe. Sonar systems are regularly employed in underwater applications and are even used 'n internal medical imaging applications. As a result of attenuation limitations, it is difficult to define an ultrasonic imaging system with an appreciable range. Secondly, propagation of ultrasonic signals is a physical molecule-to-molecule or atom-to-atom process. As such, the random thermal motion of atmospheric species is superimposed on the direction of propagation. This assures a significant beam spread with attendant loss of resolution. Also implied here is a significant problem with respect to temperature dependence.

For certain applications, ultrasonics may prove to be the technology of choice. The versatility, cost, speed, and accuracy of short range systems are highly desirable and should be explored for applications such as proximity sensing and high accuracy parts or systems inspection. A unique application of ultrasonics, phase monitoring, has recently been developed and shows great promise for specific applications (Appendix E).

With phase monitoring (PM), an ultrasonic source is directed at the object or system to be considered. The sound waves constructively and destructively interfere to produce a standing wave pattern which is sampled by an array of detectors, usually simple microphones. With a relatively small

-17-   GEO-CENTERS, INC.

computational and data storage ability, the PM system can be "trained" to recognize the pattern created by a finite data set. This ability can then be used for high tolerance automated inspection and for limited command ability.

One positive aspect of using the PM technique is that it has the limited ability to "see around corners." With optical sensing techniques, data are acquired only on surfaces that can be viewed directly. With PM, the standing wave pattern will be affected by contributions from all sources. This includes reflections (even though multiple) from surfaces not directly seen by the source.

For automated inspection, PM has proven to be extremely valuable, rapid, and accurate. An object placed in the acoustic field of an ultrasonic source will uniquely perturb the field. By sampling the field at a number of locations with an array of detectors, field deviations created by small object changes can readily be detected.

The general concept is illustrated in Figure 5. The source first illuminates a calibration object; the resultant standing field is sampled by a microphone array. The "standard" field pattern is stored in memory and the measured patterns for objects to be tested or inspected are matched against the standard. Both displacement and surface defect perturbations are detectable. Position errors as small as 1 mil (0.03 mm) and defect volumes as small as 0.002 in.^{3} (30 mm^3) have been detected at frequencies of 10 to 20 kHz.

Such sensitivity is well demonstrated by the fact that such a system can differentiate between heads or tails on a coin. Note however, that if the coin is not introduced with a consistent orientation (e.g., head always pointing in the same direction), the system loses the ability to uniquely identify status. A limited ability to accommodate rotation can be acquired by expanding the training set data base, and/or exploiting the rotational symmetry of the problem in hardware or software.

A PM system also has a limited ability to compensate for slight errors in test object placement. The source/microphone array (the relative location of source and microphones must be held fixed) may be moved and/or the object may be moved in an attempt to improve the pattern match obtained. Both theory and experiment have demonstrated, however, that if the initial placement is not

-18-   GEO-CENTERS, INC.

close to that sought, movement instructions based on sampled data may actually result in a divergence from the desired position. To prevent this from happening, initial placement should not exceed half a wavelength from the "standard" (~0.5 inch at 10 kHz).
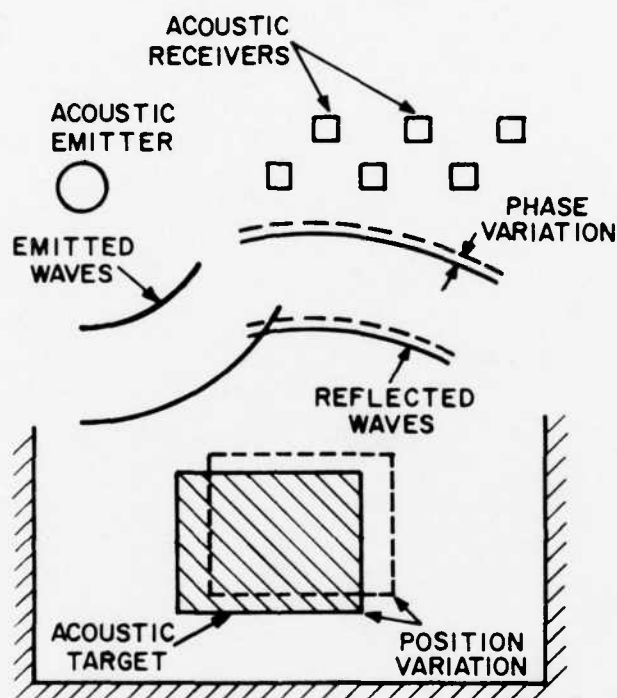
Figure 5. Outline of Ultrasonic Phase Monitoring Technique.

-19-

4. CONCLUSIONS AND RECOMMENDATIONS

The ability to acquire spatial information for robotic applications has improved considerably in the last several years. Improvements have resulted from the utilization of new technologies and from advancements in the application of older technologies. It is certain that this growth will continue and that commercial applications will provide a significant impetus to this growth. During this review, several technological areas were identified which are key to a continued growth over the long term. The following areas are specifically recommended for additional research:

• Image Correlation Techniques — needed to ensure that stereoscopy can be used in a timely and efficient manner. Recommended specifically is an examination of the use of color as a discriminant for correlation algorithms. Grey scale has been used extensively, but it is felt that the added use of color would simplify correlation algorithms and require fewer digitization levels.

Independent of the image correlation technique(s) ultimately used for stereoscopy, efforts should be undertaken to shift the image processing from the software world where it is usually developed to an implementation in hardware. This would minimize the amount of digital information which must be manipulated and would also significantly enhance data processing rates. A study of generic image processing techniques should be undertaken to determine which are capable of formulation as "on-chip" processing elements. Key to such a study would be determination of the amount of data which must be passed from pixel to pixel in an image as well as between images.

• Laser Technology — represents one of the keys to several of the data acquisition techniques reviewed in this study. To enhance this technology, two areas must be addressed:

-20-   GEO-CENTERS, INC.

1) <u>Small high-power semiconductor lasers</u> should be developed (preferably without cooling requirements). This permits the mounting of an active laser probe at positions of optimum use. It also minimizes the use of sophisticated beam transmission techniques which increase computational complexity and are difficult to maintain.

2) <u>Nonmechanical beam deflection</u> techniques need to be developed. The rotating or oscillating techniques currently used are not rugged and require considerable skill and competence to maintain alignment. Desired here would be the development of techniques similar to the acousto-optic deflectors currently used for laser printing and optical character recognition schemes. These systems are only capable of total laser beam deflections on the order of a fraction of a degree. For use in robotic applications, deflections on the order of tens of degrees are required.

● <u>Proximity Sensing</u> — needs development as a complement to the data acquisition techniques reviewed. It is recognized that all of these techniques have limited resolution abilities and that all will eventually be detrimentally affected by manipulators or other hardware as a close approach is attempted. Both Fiber Optics Systems, with a transmitted light beam, and Ultrasonics should be examined for this application. Both have shown promise and both may eventually be useful for specific applications.

In summary, recommendations are made that additional efforts be undertaken in:

- ● Image Correlation Techniques
  - Color as a discriminant
  - "On-chip" processing
- ● Laser Technology
  - Development of higher power semiconductor sources
  - Development of nonmechanical scanning techniques
- ● Proximity Sensing
  - Fiber Optics
  - Ultrasonics

GEO-CENTERS, INC.

APPENDIX A
SOLID STATE IMAGING TECHNOLOGY

The development of high-speed, accurate triangulation techniques for acquiring range information requires high resolution, solid state image sensors. The development of this technology has been rapid, but the major driving force for future progress will come from the commercial sector. Originally, solid state imaging concepts were explored for applications in space systems and in weapons or weapons delivery systems. Both applications require compact, lightweight, rugged sensors.

The commercial sector is now the major user of the technology and the attached article describing a commercial application supports this view. The capabilities reviewed here are impressive and there are indications that improvements can be expected. Effective commercialization of the concept described in this Appendix for a mass market requires high-volume, low-cost production. These benefits are of interest for robotics applications.

GEO-CENTERS, INC.

Photography
Joins the Electronic Age

JON WEINER

New low-cost computers have converted the filing process into a disappearing act for documents—that is, all documents except photographs. Data can be handled without recourse to paper, but although the technology for converting images into digital signals has been developed (witness the dramatic photographs of the outer planets), most companies cannot justify the expense for routine filing. They therefore resort to the time-honored method of retrieving dog-eared photographs from manila folders.

Not for long. In the fall of 1981 the Sony Corporation unveiled in the United States the prototype of a "filmless camera" that substitutes sophisticated electronics for film and uses a television screen instead of coated paper to display each color still shot. When the Mavica—short for magnetic video camera—is ready for distribution in 1983, it will no doubt be challenged by a host of Sony's competitors, who, despite disclaimers, are working on similar products.

Sony chairman Akio Morita called the Mavica "a revolution in photographic history," and the world press played it up as the first giant step in photography since Daguerre invented the first practical photographic process 140 years ago. But the filmless camera is really only an extension of video technology. No one expects it to replace the 35mm camera, much less drive Polaroid out of business, but the video camera will certainly find a ready market with upscale consumers who enjoy taking family snapshots. That a picture can be immediately seen on any color television—a sort of instant TV Polaroid—should be enormously appealing to amateur shutterbugs. The Mavica is also likely to find

JON WEINER is senior editor of The Sciences
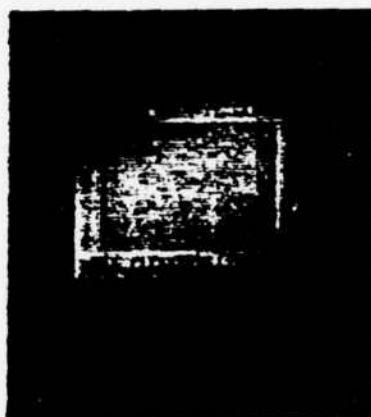
A charge-coupled device (CCD) is the heart of filmless video cameras.

a significant market in businesses that must keep a large number of pictures on file. "It will be good for news applications and other specialized applications, not for appreciating a great subject with depth," explains James Chung, who follows the photography market for Merrill Lynch.

Like the Walkman portable stereo cassette player or the Tummy TV, the Mavica bears Sony's trademark of practicality and convenience, combined in a package so miniature it invites use. It resembles a conventional single-lens reflex (SLR) camera, although it is a bit heavier at 800 grams. Like most SLR's, the Mavica has interchangeable lenses (so far, Sony plans a 25mm F 2, a 50mm F 1.4 and a 4 times zoom F 1.4 from 16mm to 64mm) and a hinged mirror to permit through-the-lens viewing. It can shoot single frames at shutter speeds from 1/60 to 1/1,000 second or make continuous recordings of up to 10 pictures per second. It shoots color pictures at ASA 200, about the speed of fast color films.

Instead of a roll of film, however, the Mavica uses a 6-by-.03-centimeter floppy magnetic disc called the Mavipak, which stores up to 50 color pictures. Essentially a small video disc, it can be inserted in a special viewer for displaying the images on an ordinary television screen.

The disc can be taken out of the camera and viewed after only a few pictures are shot and then returned to the camera. It can be erased and used over and over again, like video tape. And individual frames can easily be transferred onto video tape to make a video album. Sony has plans for a picture printer that will make color prints (five by seven inches or smaller).

Morita estimates the camera's retail price will be $650, plus about $220 for the TV-display viewer and at least $200 for the hard-copy printer. Thus the system will probably cost just over $1,000 when it first enters the market. Only the "film" is cheap: the reusable magnetic disc, in a hard plastic case, will cost $2.50.

The Mavica can be made so trim because Sony replaced the conventional vidicon tube, which is heavy and fragile, with a silicon chip that has a light-sensitive surface. This remarkable new image sensor—called a charge-coupled device (CCD)—is about the size of a postage stamp, but it accounts for a considerable part of the Mavica's price.

CCD's were invented by Willard Boyle and George Smith at Bell Laboratories in 1969. Some black-and-white miniature video cameras were made with CCD's as early as 1971. Sony has not revealed how it produces a color picture using CCD's. A CCD is a microscopic grid made up of light-sensitive squares, each of which converts the light that strikes it into an electric charge. Each of the squares represents one bit of information—a pixel, or picture element—and is approximately the size of the black dots

A-2

that make up newspaper pictures.

To transfer all these pixels into the memory of the magnetic disc, the CCD uses an electric field to pass charges to the edge of the grid. At the moment each charge reaches the edge it is measured, and the information is stored in the video disc.

At present a picture taken with the Mavica is slightly fuzzy because the CCD contains only 570 horizontal imaging elements and 490 vertical imaging elements—fewer than 280,000 pixels in all. Morita says that the resolution will improve (and expects costs to come down), but it may be many years before the CCD can match the high quality of 35mm films, whose fine grain is the equivalent of one million pixels per picture.

The beauty of the Mavica's magnetic memory is that information from it can be converted instantly into a digital signal and transmitted quickly and simply over telephone wires. A photographer halfway around the world could put the disc into a transmitter that digitizes the signals, and off the images would go to the home office. Of course, Wirephoto is nothing new to AP and UPI, but the wire services are currently forced to rely on film that is processed and printed on-site and on expensive, elaborate scanning

systems that convert the images into electronic signals.

F. W. Lyon, vice-president for news pictures at UPI, is "very interested" in the Mavica, but he has challenged Sony to improve its resolution to that of 35mm film. On the other hand, Bob Gerson, senior editor of *Television Digest*, says that if it were possible to use some of the image-enhancement techniques developed by NASA, which blend scan lines into a continuous image, then "in theory, this could give a hard print from the Mavica a lot more quality. Not great—but you're only talking about a three-by-five-inch print."

Filmless cameras will have other specialized applications, according to Harry Machida, manager for financial corporate communications at Sony. Insurance companies require millions of low-quality photographs for their records, and photographs taken by the video camera will be easy to file, store and retrieve electronically. For the same reason, the military and police will find the video system attractive.

Because the Mavica can be connected with a special adapter directly to a home video tape recorder such as Sony's Betamax, it can be used as a live video camera. There are currently 3 million video cassette recorders in

the United States, and no one expects the recent copyright ruling, which restricted video taping, to dampen sales. One out of every five VCR owners also buys a conventional portable video camera, which costs anywhere from $500 to $1,400. Against those prices the Mavica is already competitive.

Sony faces stiff competition in the future—and may not be far ahead of the pack. Sharp Corporation of Japan has announced that it is preparing to market a similar camera that weighs 270 grams less; Sharp will distribute it in Japan in the fall of 1982. Several other Japanese electronics firms and American companies such as Texas Instruments, RCA, Kodak and Polaroid are rumored to be working on similar systems.

But this kind of competition does not worry Sony overmuch. A report by securities analyst Brenda Landry of the investment firm Morgan Stanley notes that "the company would prefer to position itself in a business with good growth potential even though there may be competitors rather than have a slow-growing area all to itself." And the growth potential is unquestionable. Says Katherine Stults of Morgan Stanley, "The video revolution is real. And the Mavica becomes one more piece in that file." □

THE FILMLESS CAMERA
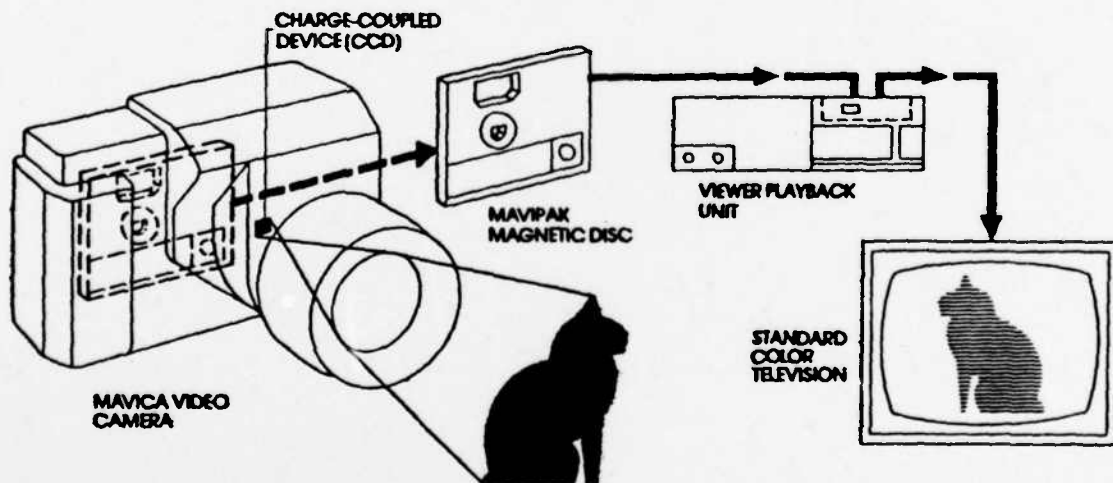
Sony's magnetic video camera uses a tiny CCD image sensor to convert light directly into electric signals. The signals are stored on a floppy magnetic disc called the Mavipak, which can store up to 50 color still pictures. The video disc is inserted into a special viewer in order to display the pictures on an ordinary television set. The disc can be erased and used over and over.

APPENDIX B
ON-CHIP IMAGE PROCESSING

Automated stereoscopy requires the development of high-speed, efficient techniques to correlate the two images to be used. Many of the concepts being explored to accomplish the desired image correlation require extensive computational effort. However, some of these mathematical processes are amenable to execution with hardware as opposed to software.

Technological advances in both active processing elements and in higher density computational elements may be capable of implementation directly on an image sensor chip. Packing densities for computational elements have steadily been increasing and have resulted in smaller, higher speed modules. Additionally, active processing has been developed which exhibits capabilities of direct interest to imaging for robotics.

A technique which can be used to acquire Fourier transforms of images on a real-time basis is presented on the following pages.* The use of surface acoustic wave technology to perform the bulk of the processing required is unique and dramatically reduces the amount of numerical processing required. Such techniques, or combinations thereof, should be explored more fully for this application.

---

*Reprinted from the July 1980 issue of Optical Spectra.

GEO-CENTERS, INC.

BEYOND OPTICAL IMAGERS: THE DEFT CAMERA

By Stephen T. Kowel

The relationship between optical imagers, such as the CCD array, and optical Fourier transformers is similar to the one between oscilloscopes and spectrum analyzers. While optical imagers make plots of image intensity as a function of position, Fourier transformers look for the spatial frequency content of optical images. Because of this difference, Fourier transformers are better suited for processing applications, including image alignment, focus detection and motion detection, than standard optical imagers.

The direct electronic Fourier transform (DEFT) sensor takes full advantage of Fourier imaging. It can electronically select arbitrary, two-dimensional, Fourier components of arbitrary images through a novel pseudo-beam steering technique.

DEFT structure

The DEFT camera (Figure 1) consists essentially of a photoconducting film of cadmium sulfide (CdS) deposited on a piezoelectric substrate (LiNbO₃). A suitable metal pattern is evaporated onto the CdS to pick up photocurrent. Interdigital transducers are used to generate two orthogonal surface acoustic waves in the substrate.

In operation, the DEFT camera focuses the optical image in its field onto the CdS film. The electric fields associated with the acoustic waves induce a nonlinear modulation of the conductivity.

A full tensor treatment of this interaction[1] reveals that the deposited contacts detect a current proportional to

$$I(t) \propto \exp{[j(\omega_1 - \omega_2)t]} \int d^2r \; I(r) \exp(-j\vec{k}\cdot\vec{r})$$

where $I(r)$ is the image intensity, ω_1 and ω_2 are the frequencies of the two acoustic waves and $\vec{k}$ has as its components the wave vectors of the two acoustic waves. By varying the acoustic frequencies, we can vary $\vec{k}$ and probe different points in the Fourier space. Under these conditions, the signal behaves as if a new acoustic wave has been created with a wave vector equal to the sum of the acoustic wave vectors. We call this effect pseudo-beam steering.

Unlike digital techniques of Fourier transformation, which digitize image information after suitable image scanning, the analog DEFT technique exploits physical material properties to extract Fourier information in real time. There is no need for digitization.

Sensing plus preprocessing

In a number of image processing applications, the unique preprocessing capability of the DEFT sensor offers advantages over alternate methods of image sensing, such as raster scanning or optical Fourier transformation.[2] Image sensing and preprocessing are combined in a single device without the need for a coherent light source, expensive optical components or precision alignment.

The utility of the DEFT technique can be illustrated by its application to some typical image processing functions. Here we outline four such examples: image alignment, focus detection, motion detection and pattern recognition.

Recently, Deft Laboratories built an experimental system for automatic image alignment with respect to a reference image using DEFT sensors. In the system, two matched DEFT sensors
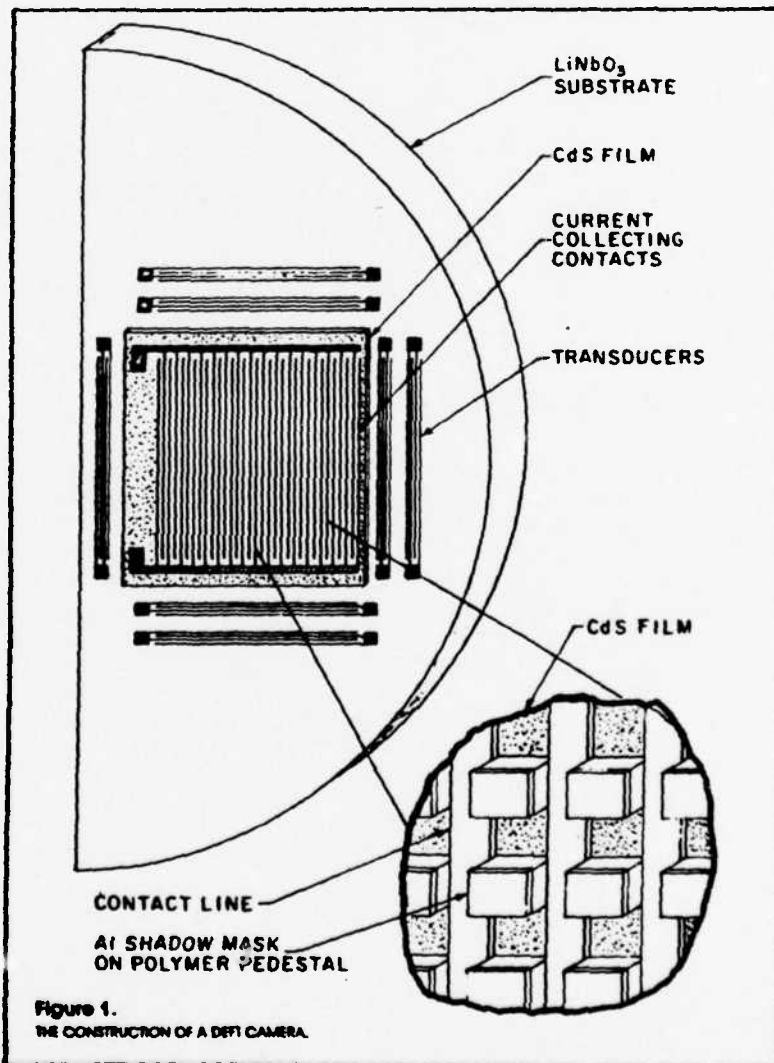

Figure 1.
THE CONSTRUCTION OF A DEFT CAMERA.

LiNbO₃ SUBSTRATE
CdS FILM
CURRENT COLLECTING CONTACTS
TRANSDUCERS
CdS FILM
CONTACT LINE
Al SHADOW MASK ON POLYMER PEDESTAL

look at two identical but misaligned images and provide Fourier components to e microcomputer. The microcomputer determines misalignment in x, y and θ, using a special algorithm based on the Fourier transform space-shifting theorem. Since the transform magnitude functions ere invariant to translational misalignment, hy computing their cross-correlation as a function of angle you can determine the angle misalignment, $\Delta\theta$. The misaligned image is then rotated to remove $\Delta\theta$.

Once $\Delta\theta$ is removed, $\omega_x \Delta x + \omega_y \Delta y$ is determined at e number of spatial fre-

quency samples using the space-shifting theorem relationship. Then Δx and Δy are determined using leest squeres estimetion.

The algorithm is compliceted by the fact thet true phase values are required but only principel velues of phase ere available. This leads to an iterative process wherehy increasingly higher spetial frequencies are used to give increasingly better estimates of Δx and Δy. The system is described in greater detail in reference 3.

The number of operations (multiplications or divisions) required to align

two imeges using the algorithm is approximetely

$$n_\theta (2n_s + 2) + n_{I_s}(5n_s + 12) \text{ops} \quad (2)$$

where n_s is the number of sample values used to represent the image or transform. n_θ is the number of angle increments used during correlation and $n_{I_s} < n_\theta$ is the number of least-squares estimetes. In contrast, alignment based on correlation of the image intensity functions requires

$$n_x n_y n_\theta (2n_s + 2) \text{ops} \quad (3)$$

where n_x and n_y are the number of x and y increments. Beceuse the DEFT sensors provide Fourier transforms. images ere not used directly and a one-dimensionel cross-correlation plus e leest-squeres estimation of two parameters ruplaces e less efficient three-dimensionel cross-correlation.

The present alignment system cen align e 25-centimeter-square image to within 50 micrometers in x and y coordinates and to 0.1 degree in θ. The next generetion of DEFT cameras will have better signal-to-noise retio and lerger bandwidth for significently better overall performence.

Automatic focus detection

Focus detection with the DEFT sensor relies on the fect thet most focused
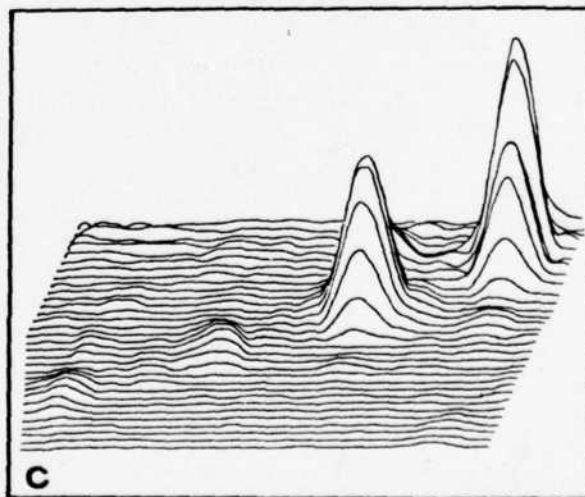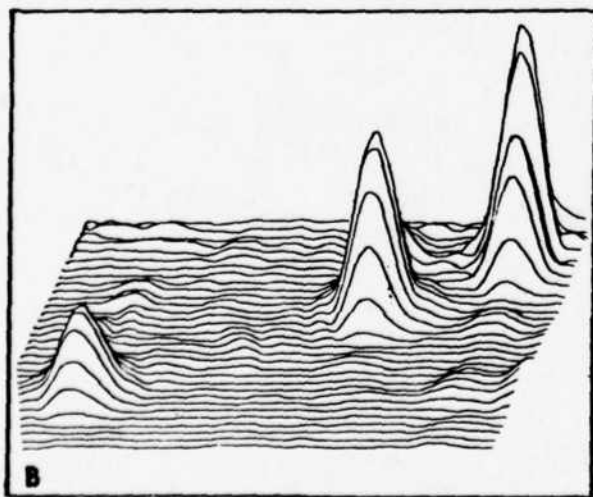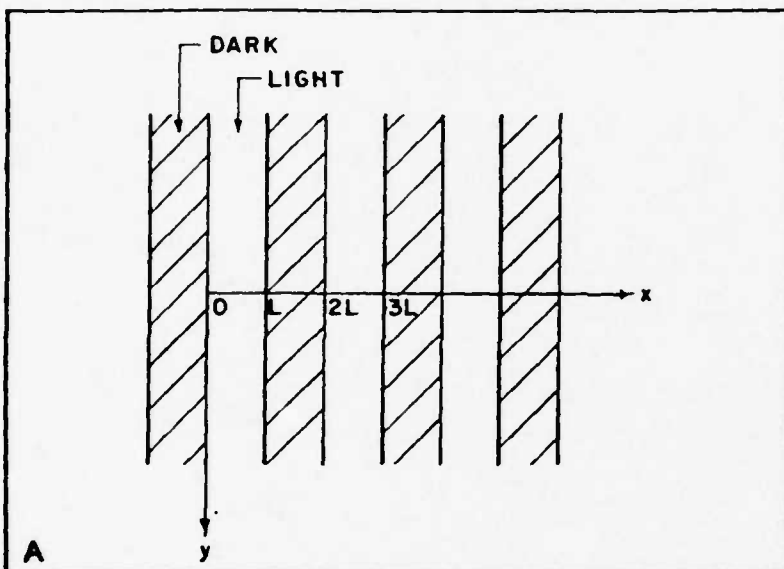
Figure 2.
THE OBJECT SHOWN IN A WAS IMAGED WITH A DEFT SENSOR B SHOWS SENSOR OUTPUT WHEN THE IMAGE OF A IS IN FOCUS C SHOWS SENSOR OUTPUT FOR AN OUT-OF-FOCUS IMAGE OF A
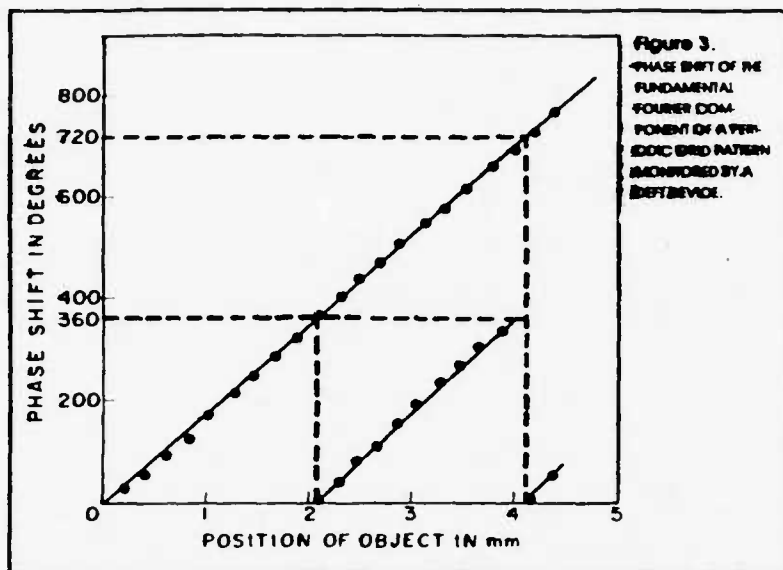
Figure 3.
PHASE SHIFT OF THE FUNDAMENTAL FOURIER COMPONENT OF A PERIODIC GRID PATTERN MONITORED BY A DEFT DEVICE.

images contain abrupt changes in intensity due to object edges. As a result, a focused image contains greater high spatial frequency content, which can be detected by the DEFT sensor.

You can visualize the sensor's potential for automatic focus detection by a simple experiment. The object of Figure 2A was imaged on a DEFT sensor. Figures 2B and 2C are plots of actual sensor output. The axes are those in two-dimensional Fourier space, and the peaks indicate the prominent spatial frequencies contained in the image of Figure 2A.

When the image is in focus, the sensor produces the output shown in Figure 2B. The largest peak is the zero spatial frequency component. The other two peaks are the first and third harmonics of the spatial periodicity in the image. Figure 2C shows the sensor output with the image slightly out of focus. Notice that the first harmonic is smaller and the third harmonic has virtually disappeared.

An automatic focus detection system using the DEFT sensor would operate as follows: The sensor views the image to be focused and the lens is moved through its extremes of travel. At a sufficiently high sampling rate, the sensor is addressed at one or more spatial frequencies. For example, let the lens be at position x_i at sample t_j. At each x_i, the magnitudes of the addressed Fourier components are read and averaged to provide a signal to be maximized. Finally, the lens is moved to the position where the maximum signal was detected.

In practice, because of the low pass nature of imaging systems and the fact

that very low spatial frequencies are insensitive to focus, it is only necessary to monitor a band of spatial frequencies. The maximum sensitivity to change in focus occurs at high spatial frequencies, but the signal strength is greater at lower frequencies.

Reference 4 gives additional details of automatic focus detection using the DEFT sensor.

Motion detection: following Fourier phase

For motion detection, the DEFT camera utilizes the spatial shifting theorem of Fourier transforms. This theorem states that a pure translation of a picture in real space leaves the magnitude of Fourier components unchanged, but alters their phase by an amount proportional to the magnitude of translation. Therefore, by monitoring the phase of a Fourier component, you can detect any motion of the picture. Indeed, because the DEFT device is a two-dimensional Fourier transformer, you can also determine the direction of motion with it.

To test this, we projected a periodic grid pattern on the DEFT sensor (see Figure 2A) and monitored the phase of its fundamental Fourier component. The results are shown in Figure 3, which illustrates that the phase changes linearly and undergoes a 360-degree shift at 2.1 millimeters. This is exactly the periodic length of the grid image on the sensor.

We have also developed a simple algorithm, explained in reference 5, that yields the velocity, as a function of time, of an object moving against a complicated background.

The DEFT sensor can be used to simplify the detection of image features, such as a key word in a page of text or a man-made object in an aerial photograph. Detection of these features with an optical imager would require a search of the field of view for the right feature. That is, each segment of the image where the feature might appear must be correlated with a reference template.

Using the DEFT sensor, it is possible to detect the desired feature in the Fourier transform of the image rather than in the image itself, because every feature has a unique Fourier transform. The advantage in using the transform arises from the Fourier transform spatial shifting theorem. That theorem specifies that no matter where a feature is within an image, the magnitude of the Fourier transform of the feature will always be centered at the origin of the Fourier plane. Therefore, it is only necessary to correlate the feature template once at the origin.

In situations where the feature may be rotated as well as translated, the template can be rotated around the origin of the Fourier plane to determine the best match. In contrast, searching for a rotated and translated feature in the image directly requires rotation of a template around the center of every segment of the image where the feature might appear. □

References

1. Kowel, S.T., et al. (1976). Two-dimensional transform (DEFT) devices: analysis, fabrication and evaluation. Report DAAG53-76-CO162. Prepared for the U.S. Army Night Vision Laboratory, Fort Belvoir, Va. 22060.

2. vander Lugt, A. (1974). Coherent optical processing. PROC. IEEE 62:10.1300-19.

3. Moyer, A.L., S.T. Kowel and P.G. Kornreich (1980). An experimental image alignment system. PROC. SPIE TECH SYMP. EAST.

4. Kowel, S.T., P.G. Kornreich and O. Lewis (1976). Focus detection using direct electronic Fourier transform sensors. J APPL. PHOTOG. ENG. 2:3:113-118.

5. Kowel, S.T., P.G. Kornreich, O. Lewis and F. Kirschner (1975). Passive detection of motion traverse to the optical viewing axis. IEEE TRANS. INSTRUM AND MEAS (September 1975).

Meet the author

Dr. Stephen T. Kowel is professor of electrical and computer engineering at Syracuse University. He is also vice president and a co-founder of Deft Laboratories, Inc., in East Syracuse, N.Y. He received a PhD in EE from the University of Pennsylvania

CONTROLLED ILLUMINATION CONCEPT

The concept outlined on the following pages represents one of the more sophisticated applications of controlled illumination. The technique used to generate the illumination pattern is unique; space coding the pattern helps minimize the amount of data which must be stored and processed.

Positive aspects of this approach include:

● no mechanical scanning,
● simultaneous large area illumination, and
● with a laser source, the illuminating array can be well controlled.

Negative aspects include:

● a high-speed electronic shutter must be developed,
● as a bistatic system, obscuration/shadowing cannot be avoided, and
● large amounts of image data must be stored and processed.

The latter point is worthy of further discussion. Consider the case where the illumination array is confined to a symmetrical M x M pattern, where to simplify later processing, M is a multiple of 2. Consider the case where a 128 x 128 array is used ($M = 2^7$). Each illumination spot can then be uniquely identified with

$$N = 2(1 + \log_2 M) = 16 \text{ images} = 2^4 \text{ images}$$

Proper use of the illuminating array could reduce this by a factor of 2, but, for generality, it will be maintained here. To adequately resolve all M x M spots should require at least a factor of 4 improvement in resolution over the number of elements to be viewed. Therefore, each image would of necessity have to be composed of

GEO-CENTERS, INC.

4xM x 4xM elements (512 x 512)

or $2^2 \times 2^7 \times 2^2 \times 2^7 = 2^{18}$ pixels. With a grey level resolution of 8 bits (2^8), effective spatial mapping would then require

$$2^4 \times 2^{18} \times 2^8 = 2^{30}$$

bits of information. This is a large number of data points to store and process. If a reasonable 2 computer operations/bit of information is assumed, and the desired coordinate map will be generated in 1 second, then a computational rate of

$$\frac{2 \times 2^{30}}{1} = 2^{31}$$

operations/second, or in excess of two billion operations/second will be required. This is at least three orders of magnitude faster than systems currently available.

Discussions with one of the authors has ascertained that the problem is reduced in complexity by using binary image coding and then creating pseudo-images for subsequent processing. The operating rate would then reduce to 2^{23} operations/second, or a minimum of 10-MHz processing rate. Such rates are above the 1-MHz rate available with today's minicomputer technology, and to be able to use this concept demands the use of considerable preprocessing and hardwired computational techniques. Both are feasible with today's technology.

Effective technological utilization requires high-speed accurate coding of the illumination array. To reach practical data acquisition rates demands that the masking used for coding be accomplished electro-optically. Mechanical shuttering is not fast enough, but it is not clear that an effective electro-optic shutter can be developed. Work on developing such a shutter is in progress.

GEO-CENTERS, INC.

Laser electro-optic system for rapid three-dimensional (3-D) topographic mapping of surfaces

Martin D. Altschuler
Hospital of the University
of Pennsylvania
Department of Radiation Therapy
3400 Spruce Street (Mail Stop 522)
Philadelphia, Pennsylvania 19104

Bruce R. Altschuler
United States Air Force School of
Aerospace Medicine
Dental Investigation Service
Brooks AFB, Texas 78235

J. Taboada
United States Air Force School of
Aerospace Medicine
Laser Effects Branch
Brooks AFB, Texas 78235

CONTENTS

Abstract. A method is described for high-resolution remote three-dimensional mapping of an unknown and arbitrarily complex surface by rapidly determining the three-dimensional locations of $M \times N$ sample points on that surface. Digital three-dimensional (3-D) locations defining a surface are acquired by (1) optically transforming a single laser beam into an (expanded) array of $M \times N$ individual laser beams, (2) illuminating the surface of interest with this array of $M \times N$ (simultaneous) laser beams, (3) using a programmable electro-optic modulator to very rapidly switch on and off specified subsets of laser beams, thereby illuminating the surface of interest with a rapid sequence of mathematical patterns (space code), (4) image recording each of the mathematical patterns as they reflect off the surface using (a) a wavelength-specific optically filtered video camera positioned at a suitable perspective angulation and (b) appropriate image memory devices, (5) analyzing the stored images to obtain the 3-D locations of each of the $M \times N$ illuminated points on the surface which are visible to the camera or imaging device, and (6) determining which of the laser beams in the array do not provide reflections visible to the imaging device. Space coding of the light beams allows automatic correlation of the camera image (of the reflected spot pattern from the surface) with the projected laser beam array, thus enabling triangulation of each illuminated surface point. Whereas ordinary laser rangefinders aim and project one laser beam at a time and expect to receive one laser beam reflection (bright dot image) at a time, the present system is optical (nonmechanical and vibration-free) and can collect all the data needed for high-resolution 3-D topographic mapping (of an $M \times N$ sample of surface points) with the projection of as few as $1 + \log_2 N$ light patterns. In some applications involving a rapidly changing time-dependent environment, these $1 + \log_2 N$ patterns can be projected simultaneously in different wavelengths to allow virtually instantaneous data collection for a surface topography. The hardware and software used in determining the (x,y,z) location of each surface dot can be made highly parallel and can handle noise as well as multiple-grazing reflections of laser beams. In common with many other active rangefinder devices, the proposed method is unambiguous in determining the topography of all nonspecular, illuminated, and visible surfaces within its operating (stereo) range, is simple to set up and calibrate, requires no *a priori* knowledge of the object to be inspected, has a high signal-to-noise ratio, and is largely insensitive to material textures, paint schemes, or epidermal properties that mask surface features to inspection by passive topographic devices.

Keywords: *three-dimensional (3-D) imaging; automated replication; robot vision; topographic mapping; pattern recognition; artificial intelligence; photogrammetry; electro-optics; laser; imaging.*

Optical Engineering 20(6), 953-961 (November/December 1981)

1. INTRODUCTION

Interest in robot vision has greatly increased recently because providing a vision capability to an industrial robot would enhance its versatility and generic utility in a factory/assembly environment. For example, computer-aided manufacturing may be improved if a real-time fully-three-dimensional (3-D) computer-aided inspection system were on-line.

A standard videocamera for robot vision provides a two-dimensional image which usually contains insufficient information for a detailed three-dimensional reconstruction of an object. (This is not always a problem, however, if the objects of interest in the robot/inspection environment can be mathematically defined and/or labeled in advance.)

To obtain the additional information needed for three-dimensional mapping of objects with complex surface shapes, a scene can be analyzed passively by stereo photogrammetry or actively with rangefinders and coded illumination. Passive stereophotogrammetry generally requires a human operator to determine corresponding scene positions in different photographs,

and is therefore too slow for real-time applications. Automated passive stereophotogrammetry requires considerable analysis.[1-4] Methods of actively interrogating a scene (by applying various kinds of light to the scene) have been used in recent years. Laser rangefinders[5,6] project one beam onto the scene at any instant; thus there is no difficulty in correlating the illuminated position in the scene with its images in each of two cameras. Although this method requires as many "images" as there are sample points in the scene, very rapid sequential laser rangefinders may soon be possible.[7] Holography requires the interference of phase coherent light beams (one beam scattered off the scene and one reference beam), but the *scene* must be free of vibrations, and to extract numerical data is often difficult. Three-dimensional information has also been obtained by illuminating the scene from different directions[8-11] and by applying light grids[12,13] and light stripes.[14-17] The light stripe method appears to have been adapted recently for commercial use to create 3-D busts and sculpture.[18]

The system described in this paper analyzes a sequence of laser dot patterns which are rapidly projected onto a surface and viewed from one or more suitable perspectives. An example of a system consists of

(1) a laser beam array generator:
  (a) a single laser,
  (b) a lens and shearing plate assembly that expands and partitions the primary laser beam into a two-dimensional (usually rectangular) array of, say, M × N separate laser beams (where M and N are typically about 128).[19,20]
  (c) a spatially programmable electro-optic light-modulating device to sequence the M × N beams through several (for example) binary-encoded patterns;
(2) an optical image recorder:
  (a) one or more video cameras (with wavelength-specific optical filters), each of which captures in digital form (or transmits in analog form to an A/D converter) the image of each coded pattern as reflected from the surface and seen from the particular camera perspective,
  (b) a device to synchronize all the TV cameras with the patterns generated by the electro-optic device,
  (c) a buffer storage device to hold a sequence of images, and/or a device to rapidly transfer image data to a computer;
(3) software:
  (a) software which rapidly decodes the sequence of TV images and calculates the position (x,y,z) of the surface at each visible dot,
  (b) a software warning capability which can automatically detect inconsistent or incomplete data (e.g., from incorrectly pointed TV cameras) and can suggest corrections to the operator,
  (c) software for image processing and error detection (possibly with an extra parity bit image),
  (d) software for interpolation between surface points to obtain a continuous surface,
  (e) software for fully-three-dimensional pattern recognition, motion detection, etc., depending on application.

An array of laser beams, subsets of which can be turned on and off by an electro-optic shutter under computer control, can be perceived as an "active camera." The present paper then discusses a modified method of stereophotogrammetry using one active camera and at least one "passive camera" rather than two passive cameras as in conventional stereophotogrammetry. (See also the preliminary accounts of this system.[21,22]) More than one passive camera may be used to view the projected patterns from more than one viewing angle if the surface of interest is rough or convoluted. Systems using several active cameras (projectors of laser beam arrays each at a selected wavelength) and several passive cameras can also be used for those applications requiring both low-resolution global mapping of large surfaces and simultaneous high-resolution mapping of selected areas, or where simultaneous viewing of multiple surfaces is desired, or if the surfaces of interest are changing in time.

The active-passive camera system has several advantages over strictly passive camera systems. Industrial parts of varied composition, material reflectivity, material finish, and textural or paint combinations, can produce artifacts or ambiguities when straightforward passive video imaging is used; interpretation difficulties and boundary definition for pattern recognition may become difficult especially when convolutions occur in complex shapes. The projection of discrete beams creates unambiguous reflective peaks (bright dots) on the object surface that are largely independent of surface characteristics, and are detectable despite peak intensity variation between dots on mixed textural surfaces. Natural protective coloration of biological specimens in their natural habitat could mask passive analysis but are clearly measurable using active sensing.

The projection of a laser beam array onto a surface produces a dot pattern image in a viewing camera. With highly convoluted surfaces, the dot images may appear much less ordered than the original beam array. Space coding, however, tags each column in the laser beam array so that the beam array column of each dot seen on the surface is uniquely identified no matter how randomized the dot images may become. Thus the reflected dot pattern (passive image) can be correlated automatically with the original beam array projection pattern (active image) to permit point-by-point triangulation of the sample points on the 3-D surface.

Beam reflections (bright dots) hidden from the passive camera sensor can be "detected" as missing by the software. This is done by taking attendance, that is, by matching the M × N discrete beams projected with those beams whose reflections are imaged. Knowledge of which beams are not imaged provides the feedback information needed to reposition a (passive or active) camera during automatic topographic scanning of a 3-D object.

Various stripe projection methods can also be space coded but require somewhat more analysis to provide information than does the discrete dot (beam array) method described here.

A space code for an array of beams arranged in M rows and N columns reduces the number of images, I, necessary for correlating all light spots seen on the surface to $I = 1 + \log_2 N$ (compared with $I = M \times N$ for a laser scanner), where N is also the number of columns of the electro-optic shutter which can be individually switched. For convenience, the value of N is usually chosen to be a power of two.

2. MATHEMATICAL METHOD

When an array of laser beams illuminates a surface, at least some of the illuminated surface positions can be imaged by an image recording device (e.g., a video camera) at a suitable perspective. The passive image plane then contains a large number of bright dots each caused by a light ray connecting an illuminated position on the surface with the passive camera focus. The information in the passive image plane (that is, the projection of some of the 3-D surface spots onto the 2-D image plane) is by itself insufficient to determine the 3-D positions of the illuminated spots on the surface.

Suppose an array of laser beams diverges from some focal point or laser source and passes through a transparent active image (shutter) plane. If (1) a particular laser beam (identified by its intersection with the transparent active image plane) illuminates a spot on the surface of interest, and (2) the image of that spot can be located in the passive camera plane, then the 3-D spatial position of the surface spot can be determined (just as in stereophotogrammetry) provided the homogeneous 4 × 4 transformation matrices (containing parameters of orientation, displacement, perspectivity, scaling, etc.) are known for the active camera (laser source and shutter) and the passive camera.

The passive camera image (x°,y°) of a point (x,y,z) in the scene is given by the perspective transformation[22-24]

$$(T_{31}-T_{14}x°)x + (T_{21}-T_{24}x°)y + (T_{31}-T_{34}x°)z$$
$$+ (T_{41}-T_{44}x°) = 0$$

$$(T_{12}-T_{14}y^\circ)x + (T_{22}-T_{24}y^\circ)y + (T_{32}-T_{34}y^\circ)z$$
$$+ (T_{42}-T_{44}y^\circ) = 0 . \tag{2}$$

If the scene-to-image transformation matrix T is known, we have for each laser dot visible to the TV camera the known quantities T_{ij}, x°, y° and two equations for the three unknowns x, y, z. We need one more equation. Suppose our laser beam array passes through an electro-optic shutter, so that the intersection of a beam with the shutter plane has a unique position (u,w) in that plane. Then the laser beam array can also be described in terms of a perspective transformation

$$(L_{11}-L_{14}u)x + (L_{21}-L_{24}u)y + (L_{31}-L_{34}u)z$$
$$+ (L_{41}-L_{44}u) = 0 \tag{3}$$

$$(L_{12}-L_{14}w)x + (L_{22}-L_{24}w)y + (L_{32}-L_{34}w)z$$
$$+ (L_{42}-L_{44}w) = 0 , \tag{4}$$

where L is the scene-to-laser transformation matrix, (u,w) identifies the particular beam in the shutter plane, and (x,y,z) is (as before) the (unknown) position on the surface (in the scene) that the laser beam hits.

We apply space coding to associate with each image point (x°,y°) the value u of the corresponding laser beam. We then have the given quantities T_{ij}, L_{ij}, x°, y°, u and solve Eqs. (1), (2), (3) above for the three unknowns x, y, z provided the equations are non-singular. Thus for each image point (x°,y°) we can obtain the corresponding surface position (x,y,z).

The camera equations by themselves give two equations for three unknowns and thus determine only the ray from the scene to the camera. The first equation for the laser perspective transformation (with u given by the space code) provides the plane which intersects the ray to the camera. Clearly, a well-conditioned solution for x, y, z requires that the laser and camera parameters (in particular, the laser and camera positions and the space coding planes $u = $ const) are such that the solution rays from the camera are not nearly parallel to the mathematical planes determined by L_{ij} and $u = $ constant. Well-conditioned solutions (accurately determined positions) should be obtainable as long as the points in the scene are not extremely distant from the camera laser system (where all distances are measured relative to the camera-laser separation distance). Once we find x, y, z, we can calculate w from the last equation so that we can later determine which laser beams (u,w) have been imaged.

3. SPACE CODING

We now describe the space coding technique. Suppose we have an $M \times N$ array of laser beams (M rows, N columns) which pass through an electro-optic shutter plane u,w. Let the centroid of beam (n,m), where $1 \leq m \leq M$ and $1 \leq n \leq N$, intersect the shutter plane at some position (u_n,w_m), such that $n \cdot 1 \leq u_n/a < n$ and $m \cdot 1 \leq w_m/b < m$, where a is the distance between the midlines of the adjacent columns of the laser beam array and b is the distance between the midlines of the adjacent rows of the laser beam array. With this definition, that beam of the laser beam array which passes through the shutter plane at position (u,w) is identified with the unique integer pair

$$(n,m) = (1 + \text{flr}(u/a), 1 + \text{flr}(w/b)) , \tag{5}$$

where $\text{flr}(x) = $ largest integer contained in real number x.

We design the electro-optic shutter to have N separately controllable columns (one for each column of the laser beam array) so that if we apply an electric signal to one of N input wires, say wire n, the domain $\{(u,w): (n-1)a \leq u < na\}$ of the laser shutter will become opaque. With such a shutter we can control which beams of the laser beam array are transmitted and which are blocked, and in this way encode patterns in the array of transmitted beams. By projecting a sequence of $1 + \log_2 N$ binary laser beam patterns, we

can determine uniquely, for any dot image seen in the passive image plane, the address n of the shutter column of the corresponding laser beam.

As an example, suppose we have a 200×16 laser beam array passing through an electro-optic shutter with 16 controllable columns labeled by $1 + \text{flr}(u/a)$. We then sequentially project (and image) the following patterns: (1) the entire laser beam array, (2) the higher-numbered half of the array (columns 16 through 9 transparent; columns 8 through 1 opaque), (3) alternate quarters of the array (columns 16 through 13 and 8 through 5 transparent; all others opaque), (4) alternate eighths of the array (columns 16, 15, 12, 11, 8, 7, 4, 3 transparent; all others opaque), (5) alternate columns of the array (even columns transparent; odd columns opaque). Altogether $1 + \log_2 N$ patterns must be projected and imaged for a binary space code. In the case just illustrated, $N = 16$, so that $1 + \log_2(16) = 5$. Then for any spot imaged by the camera during the first pattern (all the beams), the value of $\text{flr}(u/a)$ of the corresponding beam in the laser beam array is given by the binary number obtained by the appearance or nonappearance of the spot in the images of subsequent patterns. For example, suppose a surface spot appears in the passive camera images of patterns 1, 2, and 4 but not in patterns 3 and 5. Then

$$\text{flr}(u/a) = 1010_2 = 10$$

(corresponding to its appearance or nonappearance in patterns 2, 3, 4, 5) so that

$$n = 1 + \text{flr}(u/a) = 11$$

shutter domain $= \{(u,w): 10 a \leq u \leq 11 a\}$.

The apparatus can be set up and calibrated so that a more precise position value u_n on the shutter plane is known for each column n of the laser beam array. Thus once the shutter column $n = 1 + \text{flr}(u/a)$ is determined, the value of u_n is known for the laser beam.

If the transformation parameters T_{ij} and L_{ij} are known, we can obtain (x,y,z) for each (x°,y°,u). Once (x,y,z) is known, w can be found with the last equation. Any beam (u,w) not visible to the camera can be found by listing all the projected beams and then eliminating from the list those that *are* visible.

In Fig. 1 we illustrate the space coding for a 6×8 array. Only four spots on the scene are visible to the camera in this example. The laser shutter column corresponding to each image point is found in Table 1 from the data of Fig. 1.

TABLE 1.

Image point	Binary number	flr(u/a)	n
A	000	0	1
B	011	3	4
C	110	6	7
D	010	2	3

The space code method, which allows us to use one passive camera and one active camera, should be compared with the case of stereophotogrammetry in which one must use two cameras and know the corresponding transformation matrices T_{ij}^a, T_{ij}^b and the corresponding image points (x_a°,y_a°), (x_b°,y_b°) for the same unknown point (x,y,z) of the scene; hence four equations are solved for the three unknowns x, y, z with inconsistencies treated by least squares.[23,24]

4. OBTAINING THE TRANSFORMATION PARAMETERS

In a laboratory the camera and laser beam projector can be fixed and the parameters T_{ij} and L_{ij} measured directly.

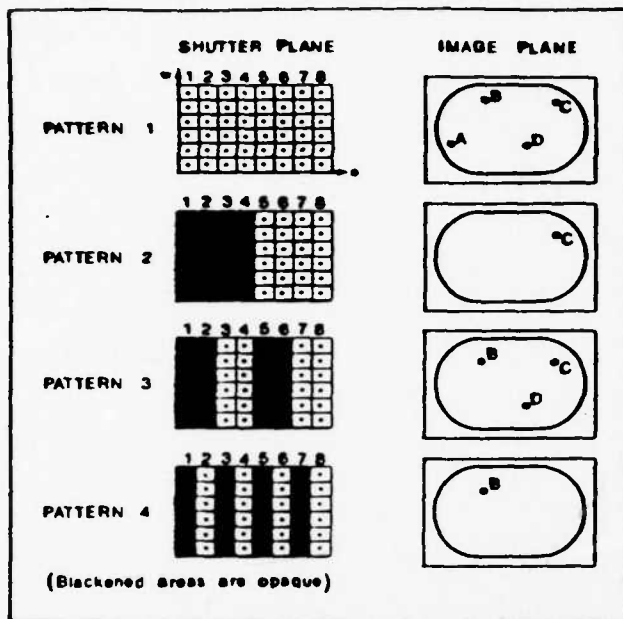In portable operation, as for example a medical or industrial ap-

SHUTTER PLANE   IMAGE PLANE

PATTERN 1

PATTERN 2

PATTERN 3

PATTERN 4

(Blackened areas are opaque)

Fig. 1. The array of laser beams is space coded by the sequence of patterns shown at the left, which is obtained by controlling the opaqueness of the columns of the shutter plane. The intersections of the laser beams with the scene provide bright spots on the scene which can be imaged from a suitable perspective by (say) a TV camera (provided there are no obstructions in the line of sight between the scene and the camera). The TV camera is synchronized with the shutter so that for each projected pattern a TV image is obtained. In the illustration, the TV image sees four spots of the scene corresponding to four laser beams. The columns of the shutter plane through which these beams passed can be determined by the space coding. Point B is seen in pattern 3 and pattern 4 but not in pattern 2 and so correlates with shutter column $1 + 011_2 = 1 + 3 = 4$.

plication, it may be necessary to set up and disassemble the equipment periodically or per object to be mapped, so a rapid calibration scheme is needed which a nonexpert can immediately use.

One simple calibration method is to insert into the scene a prism whose faces are known planes. To find the T transformation matrix, we locate six vertices of the prism in the passive image plane, so that we know the values (x_k°, y_k°) and (x_k, y_k, z_k) for $i = 1, \ldots, 6$. We can then set $T_{44} = 1$ and solve eleven equations for the eleven unknowns T_{ij} of Eqs. (1) and (2).[23,24] To obtain the L transformation matrix, we project the coded laser beam array and image the prism. If, say, $N \geq 7$ suitable points are imaged, we can obtain coefficients $L_{11}, L_{21}, L_{31}, L_{41}, L_{14}, L_{24}, L_{34}$ of Eq. (3) by setting $L_{44} = 1$ and minimizing

$$F = \sum_{k=1}^{N} \left[(L_{11}-L_{14}u_k)x_k + (L_{21}-L_{24}u_k)y_k \right.$$

$$\left. + (L_{31}-L_{34}u_k)z_k + (L_{41}-u_k) \right]^2 , \tag{6}$$

where the subscript k refers to one of the N points imaged by the passive camera. The values (x_k, y_k, z_k, u_k) can be found for each imaged point k because we know the T_{ij} (from before), the image position (x_k°, y_k°), the space code, and the planes of the prism. The least-squares solution is obtained by solving the set of seven linear equations

$$\partial F/\partial L_{ij} = 0 . \tag{7}$$

one for each of the seven unknown L_{ij} coefficients. A solution exists if the 7×7 matrix is nonsingular. In most cases, a nonsingular matrix occurs if the prism is within the active-passive stereo range and the N points appear on more than one face of the prism. In this method, the accuracies of the L_{ij} values depend on the accuracies of the T_{ij} values. This automatic calibration procedure does not provide the L_{ij} coefficients of Eq. (4), so that solving for the shutter row w after x,y,z has been obtained (that is, attendance-taking information) would not be possible. To obtain the L_{ij} of Eq. (4), a second automatic calibration procedure involving a row-coded electro-optic shutter can be easily appended.

The easiest situation of all would be to do all calculations relative to a coordinate system fixed to the laser projector or the main body of the robot and to sense independently the camera orientation and other parameters with respect to this reference frame (say by means of separate optical, mechanical, or inertial guidance systems). With such sensing capability, the projector and the sensor can be independently manipulated and need not be limited to predetermined, precalibrated viewing positions.

No matter what calibration is used, the cameras, active or passive, need not be at the same relative distance from the object of interest but can be placed at any convenient locations within the stereo range. The only restriction in positioning the active and passive cameras is that the solution of Eqs. (1), (2), (3) must exist and be well conditioned.

5. OPTICS FOR LASER BEAM ARRAY GENERATION

Laser light can be expanded with a lens and convolved into rectangular grid patterns by a parallel arrangement of fringe-generating shearing plates. However, for spatial encoding and convenient light image capture by videocamera, an array of lighted, discrete (Gaussian isolable) points is desirable. We now describe a series arrangement of shearing plates to generate the discrete (bright) beam array pattern required.

The principle used in creating a fringe array is outlined in Fig. 2.[19,20] Interference fringes are generated by front and back reflection from a shearing plate. The condition for the creation of N interference fringe lines is[19]

$$N = (2d/\lambda)[\tan^{-1}(W_o/2f_1)][\sin 2\alpha]/(n^2-\sin^2\alpha)^{1/2} , \tag{8}$$

where d is the shearing plate thickness, n is the index of refraction of the plate, W_o is the diameter of the incident laser beam, f_1 is the focal length of the focusing lens, λ is the wavelength of the laser light, and α is the incident angle of the principal ray.
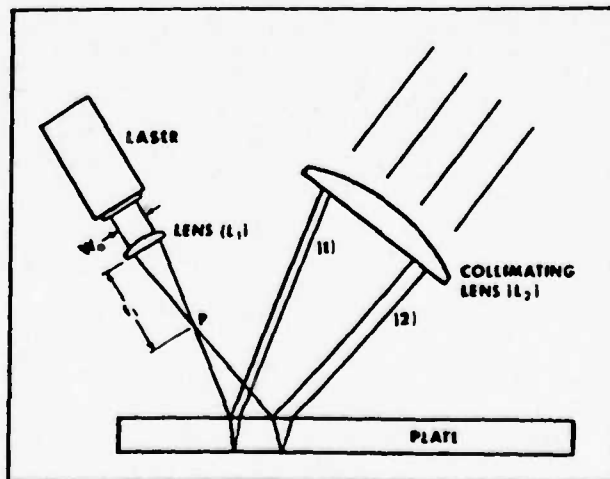

Fig. 2. Generation of parallel fringes.

If fringe lines produced by one plate are incident on a second partially silvered plate in such a way that the plane of incidence of the second plate is perpendicular to the first, the fringe lines will again be broken by interference minima, and the light reflected from the second plate will constitute an expanding rectangular array of lighted beams emanating from a common point. This light array can be recollimated and spatially filtered to produce a sharp discrete array of laser light beams. The final array may then be caused to converge for high resolution pattern recognition onto microscopic objects, or to expand at a convenient angular divergence to map larger objects. A suitable combination of lenses is used. The relative size of the beams with respect to the distances between beam centers (measured in a plane perpendicular to the array path, i.e., parallel to the shutter plane) remains constant as the total array size and individual beam diameters increase over distance in a diverging pattern.

A decided advantage of this method to generate beams in an ordered array is that the discrete beams that form the pattern are the product of diffraction, the beams behaving as individual lasers. Patterns will not deteriorate at long range (in a vacuum), and can (theoretically) be converged to wavelength separation between beams. This is in sharp contrast to forming a beam pattern by conventional masking apertures. Masks result in diffraction effects that seriously limit the useful range of the beam array and its degree of useful convergence.

A laboratory demonstration prototype of the basic hardware concept is illustrated in Fig. 3. The system generates and projects a programmed sequence of coded arrays of discrete bright laser beams onto an object or specimen. A 2''-diameter plaster stone model of a human premolar tooth (object) is seen at the lower right mounted on a lens holder. The bright beams are generated optically
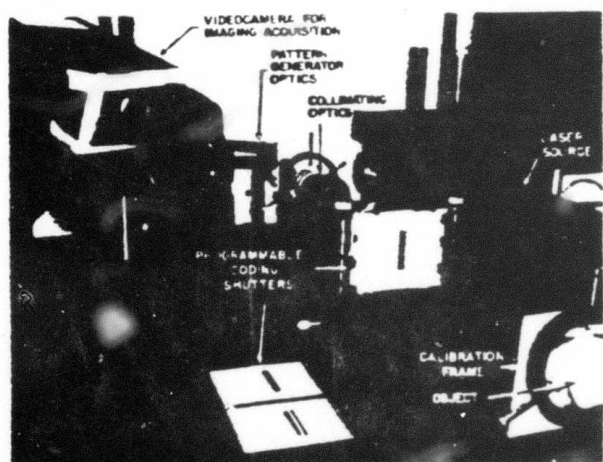

Fig. 3. Laboratory demonstration prototype of Topographic Mapping System using manually placed cards to illustrate pattern changing for space coding.

from a single linearly polarized laser source and an optical arrangement which includes beam expansion, two optical shearing plates, and beam steering devices ("pattern generating optics" of Fig. 3). Any suitable laser wavelength or set of wavelengths can be used. For this demonstration, the 514.5 nm line of a model 165 Spectra-Physics argon-ion cw laser was chosen. The optical pattern may be intermittent or continuously generated depending on the laser source, configuration, and application requirement. Multiple-source projection is possible provided that either synchronization or selective wavelength filtering is used with the scene capture imaging device; for *simultaneous* multiple source projection, the imaging device(s) must have multispectral discrimination capability.

6. HARDWARE FOR BEAM ARRAY CODING

The beam array once generated is recollimated as required by collimating optics (Fig. 3) and then passed through a programmable set of rectangular (columnar) shutters. The shutter assembly can be constructed from mechanically actuated slits, rotating assemblies of rulings, liquid crystal gates, or electro-optic or acousto-optic crystal gates. The function of the shutter is to eliminate selected columns (or rows) of beams from the laser beam array. For an electro-optic shutter, each individual shutter column (or row) becomes transmissive or opaque to the projected light beams depending on the particular electrical signal it receives. The collective effect of blocking individual columns (or rows) dictates the particular pattern being transmitted. The shutter columns (or rows) are controlled by a computer or microprocessor which drives the shutter to transmit a programmed sequence of two-dimensional array patterns onto a surface of interest. The reflected image of each new beam array pattern from the surface is recorded by one or more image acquisition devices (such as a videocamera). The shutter and the imaging devices are synchronized by a microprocessor.

The manually positioned (programmable) coding shutters shown in Fig. 3 are presented simply to illustrate the basic concept of shutter pattern reconfigurations required for space coding. A larger (128×128) beam array was masked by a white card set in a holographic plate holder to create an (illustrative) eight-column-wide pattern. The space code sequence begins when an unshuttered (uncoded) beam array is first projected on the surface of interest and is recorded by a suitable videocamera and storage electronics. In this demonstration a standard closed circuit (RS-170 compatible) General Electric model TE-26 videocamera was used (upper left corner of Fig. 3) and focused on the tooth model (seen in right foreground of Fig. 3).

After this first view is recorded the shutter is changed or reconfigured. For illustrative purposes, the second shutter configuration may be represented by the top white card lying on the optical table in Fig. 3. This shutter, when placed in the shutter holder, eliminates four consecutive columns of beams while allowing four consecutive columns of beams to pass unimpeded. The scene is again recorded to obtain the second view. The third shutter, (bottom white card on optical table, Fig. 3) replaces the previous shutter to produce a new pattern code. Figure 4 is a photograph of the video image recording of the third shutter pattern as projected on a tooth model. (Compare this camera view with pattern 3 of the shutter plane in Fig. 1.) The fourth shutter (seen in the beam array path, Fig. 3) is then placed and the scene recorded to produce the fourth view.
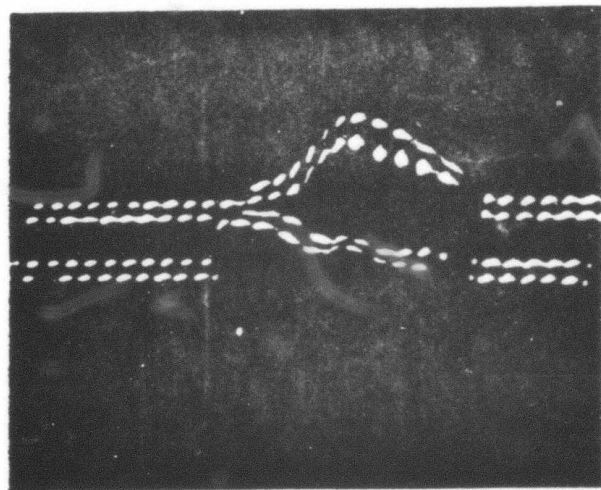

Fig. 4. An example of one shutter coded pattern as projected on a tooth model using an 8 columnar (programmable) shutter.

Manually placed shutters as illustrated (Fig. 3) in the laboratory demonstration (even if engineered into a filmstrip mechanical device) would be cumbersome and prone to failure. The optimum method for blocking selected columns (or rows) is to use a fixed, vibrationless, programmable spatial light modulator (PSLM) with exclusively electro-optic components and properties.

A laboratory demonstration model using a 128 (vertical) column electro-optic PSLM is seen in Fig. 5. The PSLM is visible in the center of the figure, just below the videocamera lens and above the top of the laser (Fig. 5). The tooth model and support is to the extreme right in the figure. In this photo, the PSLM has been programmed with four sets of "ON" (transmissive) blocks of columns alternating with four sets of "OFF" (opaque) blocks. Each set has sixteen successive columns "ON" or "OFF." A polarizing plate mounted in a metal frame appears just in front of the shutter. Not visible (hidden by the videocamera) is an interface and microprocessor which controls the PSLM.

Since space coding image conservation made a log-base-2 width system desirable and since standard U.S. videocameras of 480

Fig 5. Laboratory demonstration breadboard in which an electro-optic programmable spatial light modulator (PSLM) is substituted for manual cards (illustrated in Fig 3) to provide space coding. The PSLM blocked selected (preprogrammed) columns (or rows depending on orientation) in the projecting beam array path. The PSLM (center of photo below videocamera lens and above laser) is shown with 16 consecutive rows on, 16 off, etc., or 128 columns in which eight sets of 16 columns are alternately on (transmissive) or off (opaque or black). A microprocessor (hidden by videocamera lens) controls the PSLM pattern. Any combination of the 128 individual columns in the PSLM may be selected for transmission or opaqueness. Tooth model is at right center fixed to a magnetic mount.

useful scanned lines were available in the laboratory, a total column number of 128 was chosen for this demonstration. The PSLM module was constructed by Applied Optics Corporation (a division of ITEK Corporation) and mounted on a glass plate with a total optical (light modulating) area of 2 × 2 inches square. The module contained integral (embedded) peripheral electronics to assist in loading the pattern. Each column of electro-optic material in the shutter is individually programmable. When an appropriate voltage signal is applied to this material it will exhibit half-wave plate behavior. Since the beams in the array have definite polarization derived from the laser source, any beam passing through an energized column would be eliminated by an appropriately oriented analyzing polarizer placed beyond the shutter.

An interface was devised by Capt. Michael Tworek, USAF, so that any suitable microprocessor could be used with the PSLM. Moreover, the requirement to load the PSLM with 128 bits of

clocked serial data from the microprocessor is avoided. Instead, a 16-bit parallel-load-to-serial-data-out capability was created where eight 16-bit words are loaded into the interface to properly code the 128 columns of a selected configuration. The interface matched an available Intel SBC 80/20 single-board computer to the PSLM liquid crystal materials rise time of 500 ms. Were a faster PSLM to be used, the existing interface could be redesigned with a faster clock to load a pattern in 1 ms. Two PSLM modules and interfaces were assembled in the optical path so that either vertical columns or horizontal rows, or a combination, could be space coded. Since the present PSLM module is too slow for switching between video raster scans, the preprogrammed patterns were loaded into the microprocessor and a manual binary switching method was used to reconfigure the shutter. This was sufficient to demonstrate space coding by a PSLM in an active camera. A second generation PSLM can be constructed to permit reconfiguring the PSLM columns within (and synchronized to) the vertical blanking time of a standard video raster scan, or during the erase/recovery cycle of a solid-state imaging camera. A 100 μs reconfiguration rate could be achieved with little difficulty (well within the 1.4 ms of the vertical blanking time) using PLZT material, and even 1 μs reconfiguration rates are theoretically possible for some candidate methods.[25, 26] The limiting factors would then be the videocamera image acquisition and throughput/erase/recovery cycle, and the ability of a suitable device to store or process incoming data.

7. BEAM ARRAY PROJECTION ONTO THE SCENE

The choice of optics required to adjust the total area illuminated by the beam array, as patterned by the coding subsystem, is determined by the relative size and distance of the object from the beam array projector, and the resolution required for mapping the scene or surface of interest. Either a nondiverging, converging, or diverging beam array may be used to suit a particular application. A converging optical arrangement may be used to project a higher spatial frequency array onto a very small object. For example, a 10 dot/mm spatial frequency array (100 μm dot separation) has been projected onto a surface and viewed through a microscope objective. A zoom lens may quickly and conveniently alter the spatial frequency and size of the beam array projected field for multipurpose uses, or for changing from wide area to detailed scanning. Suitable rotatable and gimbal-mounted optical devices may be used to accurately control the direction of beam array projection. The laser beams in the array are focused at infinity by the beam array generating optics. The laser beam array will project through space for a considerable distance until impeded by an object or surface.

8. IMAGE ACQUISITION

A videocamera records from a suitable perspective the images of the sequence of coded beam array patterns reflected off the surface of interest. In Fig. 6, a full (128 × 128) pattern is intercepted by a 2"-diameter plaster premolar model which was bordered by black rubber. A General Electric TE26 vidicon camera with a GE f/2.5 to 22 and 20 to 80 mm tele-zoom was used to record the pattern. Illumination of the object was approximately 10 μW/cm^2 by 514.5 nm laser light. Ambient daylight fluorescent lighting was present. A ± 10λ selective filter was used in front of the telephoto lens. The image was recorded on an IVC 1" reel-to-reel video recorder and played back on a black and white 19" monitor. A 35 mm still camera recorded this photograph (Fig. 6) during playback. A dot in the lower left corner (beyond the model) is an artifact in the monitor. Approximately three thousand sample points are visible.

Figure 7 is photographed in real time in the laboratory. Note ambient lighting levels seen in Figs. 5 and 7. A two "ON," two "OFF" alternating pattern was being projected.

Figure 8 is another image of the object shown in Figs. 6 and 7. In this image the pattern was configured with alternating eight columns on and eight columns off. Variations in dot brightness (intensity) and adjacent dot to dot spacing are caused by surface curvatures of the object.
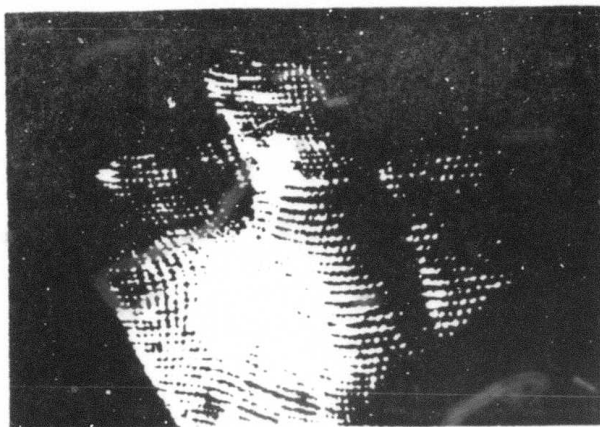
Fig. 6. Two-inch-diameter bicuspid tooth plaster model illuminated by full beam array pattern (128 × 128). Approximately 3000 dot reflections are visible. Photo was taken of video monitor while videorecording playback. Object was illuminated with a 514.5 nm argon laser in a daylight-fluorescent lighted laboratory. Videocamera had a selective-wavelength filter in front of its lens to remove ambient lighting (black cylinder housing on videocamera lens is seen in top left of Fig. 5).
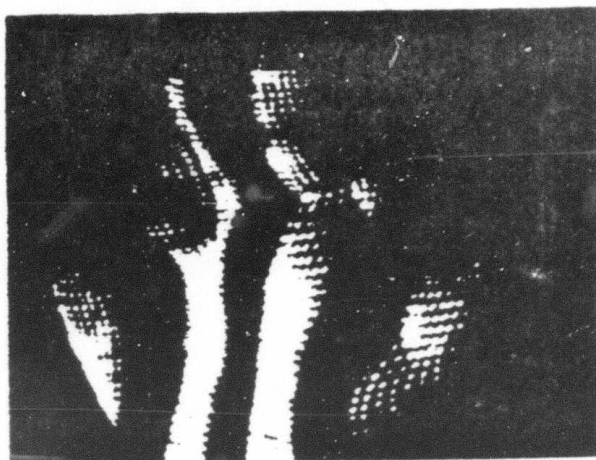

Fig. 8. Photograph of same object as in Fig. 6 but with the beam array now coded by the PSLM into one of the binary patterns. A total of eight different binary patterns are sequentially projected to completely code a 128 × 128 beam array and tag each array column for identification in the image.
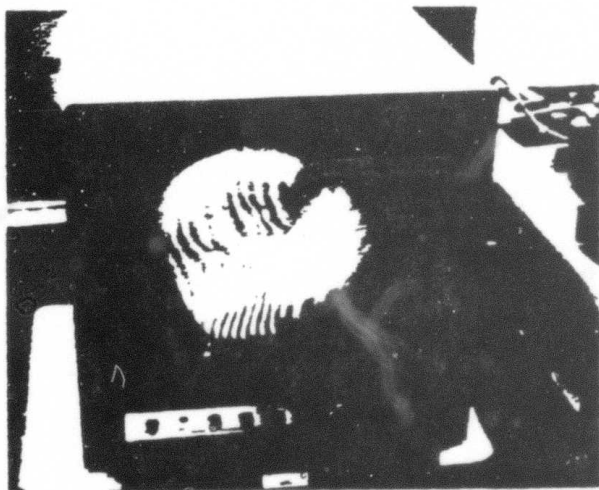

Fig. 7. Photograph of videomonitor viewing in real time a two-on-two-off consecutive pattern projected onto the tooth model. Ambient lighting in laboratory is used for this exposure.
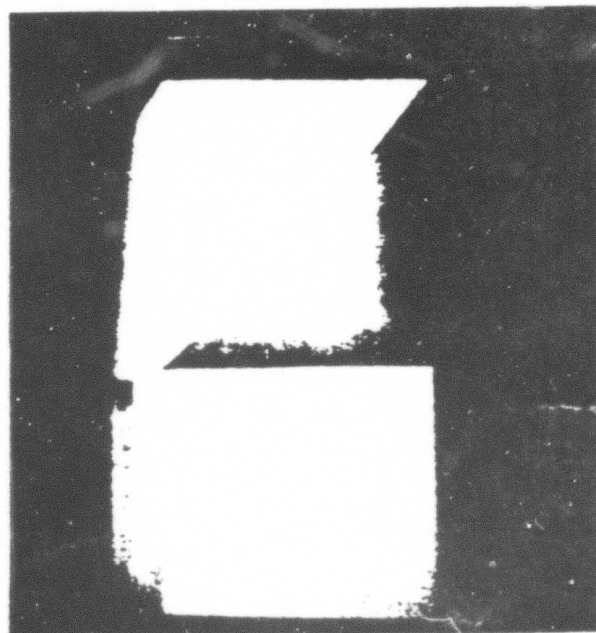

Fig. 9. White-light digitized image of prism block displayed on a computer graphics terminal after playback from digital (computer) tape.

Figures 9 and 10 show digitized images of a prism played back from computer tape onto a Comtal display unit. Figure 10 shows the passive image of a laser beam array pattern coded by a PSLM and reflected off the prism.

For a standard composite video signal, a raster scan of each image is completed in 1/30th second. During the interval between the end of the preceding raster scan and the beginning of a new scan, the electro-optic shutter reconfigures. A microprocessor, minicomputer, or clock synchronizes the videocamera, beam array shutter, and video-recording device. A video recorder or video disk system may be used to initially record the images. An alternative method would be solid-state storage of individual video images on charge-coupled device (CCD) or random access memory (RAM) chips with high speed sequential or parallel analog to digital (A/D) conversion. A 512 × 512 × 8 pixel array from a 1/30th second raster scan

may be stored in digital form (using commercially available devices) with each point mutually addressable.

For a 16,000 point topographic spatial map, a total of eight coded images is required. Total storage time is thus 8/30th second (= 267 ms), and eight images must be stored for processing. While adequate for many applications, faster speeds are often desirable. If charge-injection device (CID) or CCD direct imaging technologies are used, then the speed of total scene capture (that is, eight views for a 16,000 point 3-D) might be extremely rapid. A two millisecond time cycle is reasonable for acquisition of all data need-
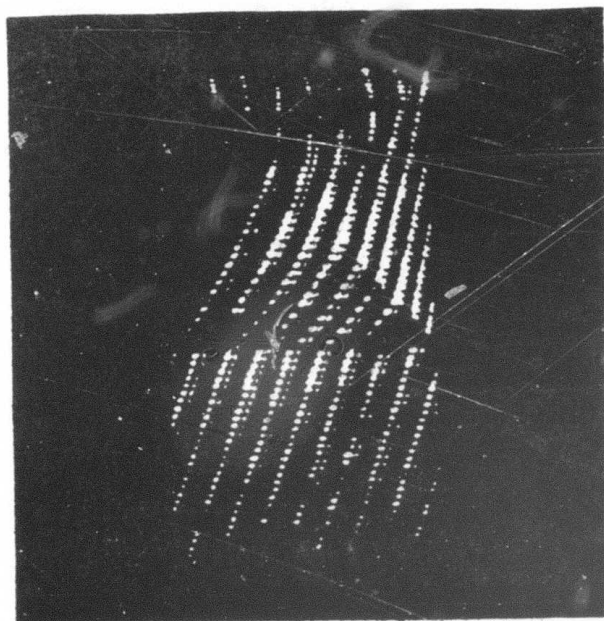
Fig. 10. Digitized image of the same prism block (displayed on a computer graphics terminal) with the 128 × 128 laser beam array shown coded by a pattern of the 128-column PSLM.

ed to define the three-dimensional coordinates (x,y,z) of a beam array projection such as the one on the bicuspid tooth model shown in Fig. 6. A 512 × 512 CID or CCD array matrix would have to be developed for rapid acquisition of data, in which the output would entail simultaneous (parallel) output of rows or blocks of rows of the array (rather than using the RS-170 output rate serial-scan methods presently used to emulate the vidicon camera raster scan). Present commercial devices have designed the imaging array output serially only, unfortunately. A parallel-processing computer, or a hierarchy of serial-processing computers would be needed to absorb data rapidly.

In certain applications pertaining to a rapidly changing environment (e.g., assembly lines, human motion, vision aids for the blind,[29] etc.), all the necessary coded images could be obtained by using a cluster of $1 + \log_2 N$ separate active cameras, each projecting a single beam array pattern at a different laser wavelength and a cluster of $1 + \log_2 N$ separate selectively filtered passive cameras, each receiving the corresponding pattern reflections in the corresponding wavelength. Such a $(1 + \log_2 N)$-fold parallel active-passive camera system would accomplish data collection for an entire topography in one cycle time. The calibration for the cluster cameras would be only slightly more complicated than for the sequential cameras.

Parallel processing of data could result in "almost instantaneous" scene capture and analysis; an updated 3-D map could possibly be calculated every second, or faster. A two millisecond scene capture with one to ten seconds minicomputer processing would be adequate for many medical and industrial applications. A high-speed synchronized cinematic process could be used as an alternative for image acquisition, but it would require a possibly unacceptable postprocessing film development delay.

One or more scene capture devices (e.g., videocameras) may be placed around the scene which is being illuminated by the space coded beam arrays. Gimbal mounts may be used for these cameras, as well as various robot arm mountings, and various electro-optic and/or acoustic aimers can be used to keep the cameras focused and centered on the scene, even if the camera(s) and beam array

projector are mobile. Calibration and positioning requirements have already been discussed. Camera-projector angulations can be calculated automatically by the minicomputer dedicated to the 3-D analysis. Suitable lens systems, determined by the user for each application, are mounted on the imaging cameras.

User-defined narrow-bandpass optical filters for each relevant laser wavelength are chosen for viewing in ambient lighting conditions. If eye safety constraints require low levels of laser light for measuring human subjects,[29] image intensification may be required. A beam array pattern illuminating at $1 \mu W/cm^2$ at 514.5 nm has been detected reflecting off a coarse black cardboard when appropriately selected videocameras or externally mounted image intensifiers were used.

9. FURTHER DATA PROCESSING AND OUTPUT

Data processing is required for (1) hardware control to synchronize the programmable laser-shutter encoder with the image acquisition system, and (2) the calculation of the three-dimensional number maps for the surface topography. The data processor may also verify the accuracy of the camera orientation vis-à-vis the laser surface target and warn of reflected points hidden to the camera by convolutions in the surface topography. Data output may be in the form of digitally stored numerical tables. Numerous applications and postprocessing methods are possible once a numerical table of the 3-D coordinates of sufficient sample points of a surface is available. For example, graphic display output and numerical control of machines are possible. Serial maps may be rapidly updated to discern movements and distortions occurring on a surface being scanned by the topographic system. Straightforward image subtraction techniques using sequentially generated output maps could lead to notification and 3-D vector analysis, in real time, of objects entering or moving within even the most cluttered scene area.

10. DISCUSSION

The objectives of this effort were to define and conceptually create the technological capability: (1) to obtain unambiguously and from a distance the three-dimensional coordinates of a large number of sample points of any open uncovered surface of any object (of any size) and thereby to determine the surface topography of the object, (2) to map a surface rapidly, remotely, nondestructively, and automatically, (3) to obtain, where possible, spatial data from as few as one scene-viewing direction, (4) to generate the topographic data in real time, (5) to design a simple-to-operate, self-contained, portable system whose output requires no human processing and whose setup requires no tedious calibration or excessive expertise, (6) to design a device so that its software determines insofar as possible the geometric transformations among the sensor, emitter, and scene (to enable decoding of the image data), (7) to design a system for maximum versatility and ruggedness for use in hazardous industrial, clean room clinical and surgical, and military environmental conditions, (8) to provide human or machine interactive capability for design fabrication, pattern recognition and scene analysis in three dimensions, and servo-controlled robot-vision-coupled applications.

The specific concept described and demonstrated by laboratory breadboarded means, and outlined mathematically, has certain uniquely interesting features that lend themselves to industrial and medical applications, were the concept developed further.

(1) The same apparatus could map microscopic or large-size objects merely by choosing suitable end objective lenses. The beam array is composed of discrete beams at infinite focus, so that an object intercepting a beam need not be at a predetermined distance from the origin of the array. Pattern array generation is the product of diffraction, so that the array may be enlarged or reduced at will or projected a long distance, avoiding the diffraction-limiting effects that occur when "grids" are produced with limiting apertures or masking slits. No moving parts or mechanical devices are used for array generation or for space coding, and the electro-optic can-

didate materials appear to have long duty cycles, possibly permitting the active camera optics to be sealed for maintenance-free use. With reasonable selection of optics, the array generation would work with most laser wavelengths or power outputs.

(2) The active (projector) and passive camera(s) can be positioned (moved) independently at arbitrary axes (e.g., on independent robot arms) for observation of the scene from different perspective directions after each set of data is acquired. Thus a fixed rotation axis for the camera system and/or the part to be inspected is not necessary.

(3) The active illumination of medical or industrial parts eliminates ambiguities obtained with passive systems due to (a) natural protective coloration, and (b) variations in reflective properties as a result of mixed textures, machine surface finishings, differing materials, coatings, and clothings. Active beam illumination creates discrete peak reflections (even if of variable amplitude) on the object surface, and a known finite array of ordered beams can be accounted for as either present or missing.

(4) Any suitable recording sensor(s) could be used for the passive camera(s). A videocamera output by whatever means could be rapidly converted to a standard TV broadcast system for remote transmission of image data in very near real time, permitting remote monitoring of automated system 3-D measurement. With lasers of different wavelengths and correspondingly filtered videocameras, more than one beam array pattern may be projected and simultaneously and independently viewed. Thus simultaneous coarse/fine scanning and global/regional scanning of a large object or parts (for positioning or comparative dimensional analysis) is possible.

(5) Gross deformations may be comparatively and accurately measured against a data base. Fully-3-D deformation analysis in gross increments is especially useful for comparison of rough castings or forgings, measurements of bubble expansion, and choosing which of a set of similar large industrial assemblies should be mated to another with the least hand fitting.

11. ACKNOWLEDGMENTS

12. REFERENCES

1. Sobel, J., 3rd Int. Joint Conf. on Artificial Intelligence, 648 (1973).
2. Shapira, R., 5th Int. Joint Conf. on Artificial Intelligence, 628 (1977).
3. Gennery, D. B., 5th Int. Joint Conf. on Artificial Intelligence, 576 (1977).
4. Moravec, H. P., CMU-RI-TR-3 (Carnegie-Mellon Robotics Institute), (1980).
5. Dreyfus, M. G., Dreyfus-Pellman Corp., Contourmate 7000 System (1979).
6. Lewis, R. A., and Johnston, A. R., 5th Int. Joint Conf. on Artificial Intelligence, 762 (1977).
7. Kanede, T., Proc. SPIE 283, 48 (1981).
8. Horn, B. K. P., MAC TR-79, MIT Project MAC (Nov. 1970)
9. Horn, B. K. P., Obtaining Shape from Shading Information, in *The Psychology of Computer Vision* (P. H. Winston, ed.), McGraw-Hill, New York (1975).
10. Shirai, Y., and Tsuji, S., 2nd Int. Joint Conf. on Artificial Intelligence, 71 (1971).
11. Woodham, R. J., 5th Int. Joint Conf. on Artificial Intelligence, 635 (1977).
12. Will, P. M., and Pennington, K. S., 2nd Int. Joint Conf. on Artificial Intelligence, 66 (1971).
13. Rocker, F., and Kiessling, A., 4th Int. Joint Conf. on Artificial Intelligence, 669 (1975).
14. Sirai, Y., and Suwa, M., 2nd Int. Joint Conf. on Artificial Intelligence, 80 (1971).
15. Agin, G. J., and Binford, T. O., 3rd Int. Joint Conf. on Artificial Intelligence, 629 (1973).
16. Popplestone, R. J., Brown, C. M., Ambler, A. P., and Crawford, G. F., 4th Int. Joint Conf. on Artificial Intelligence, 664 (1975).
17. Lovesat, E. J., Proc. Symp. Comm. V, ISP Biostereometrics, 74 311, 147 (1974).
18. Solid Photography advertisement with photograph of subject illuminated by light stripes.
19. Taboada, J., and Altschuler, B. R., Appl. Optics, 15, 597 (1976).
20. Altschuler, B. R., Taboada, J., and Segreto, V. A., Proc. SPIE 70, 347 (1975).
21. Altschuler, B. R., Taboada, J., and Altschuler, M. D., Proc. SPIE 182, 192 (1979).
22. Altschuler, M. D., Altschuler, B. R., and Taboada, J., Proc. SPIE 182, 187 (1979).
23. Sutherland, I. E., Proc. IEEE, 62 (2), 453 (1974).
24. Rogers, D. F., and Adams, J. A., *Mathematical Elements for Computer Graphics*, McGraw-Hill, New York (1976).
25. Land, C. E., Opt. Eng. 17 (4), 317 (1978).
26. Oliver, D. S., Opt. Eng. 17 (3), 288 (1978).
27. Blaha, W. P., Lipton, L. T., Wiener-Avnear, L., Grinberg, J., Reif, P. G., Casasent, D., Brown, H. B., and Markevitch, B. V., Opt. Eng. 17 (4), 371 (1978).
28. Cutchen, J. T., Harris, Jr., J. O., and Laguna, G. T., Sandia Laboratories Tech. Report SLA-73-0777 (1973).
29. Altschuler, M. D., Posdamer, J. L., Frieder, G., Manthey, M. J., Altschuler, B. R., and Taboada, J., Proc. Int. Conf. on Cybernetics and Society, ISSN 0360-8913, Cambridge, Massachusetts, p. 1000, Oct. 1980.

APPENDIX D
LASER SCANNER CONCEPT

The following pages outline a laser scanner system (Figure D-1) and its operational parameters. Although this particular system was designed by ERIM, similar systems have been assembled and tested by JPL, Perkin-Elmer, and others. Implementation details vary, but all have essentially the same basic configuration.

Positive attributes of this concept include:

- no lost data due to obscuration or shadowing,
- computational and image storage requirements are minimized,
- in the CW mode, reflectance data are also obtainable for imaging, and
- system may be directed to interrogate one point of interest without interrogating the entire field.

Areas which could use additional work include:

- higher power laser sources, particularly semiconductor,
- more efficient, nonmechanical means to achieve beam steering,
- inherent range ambiguity in CW mode, and
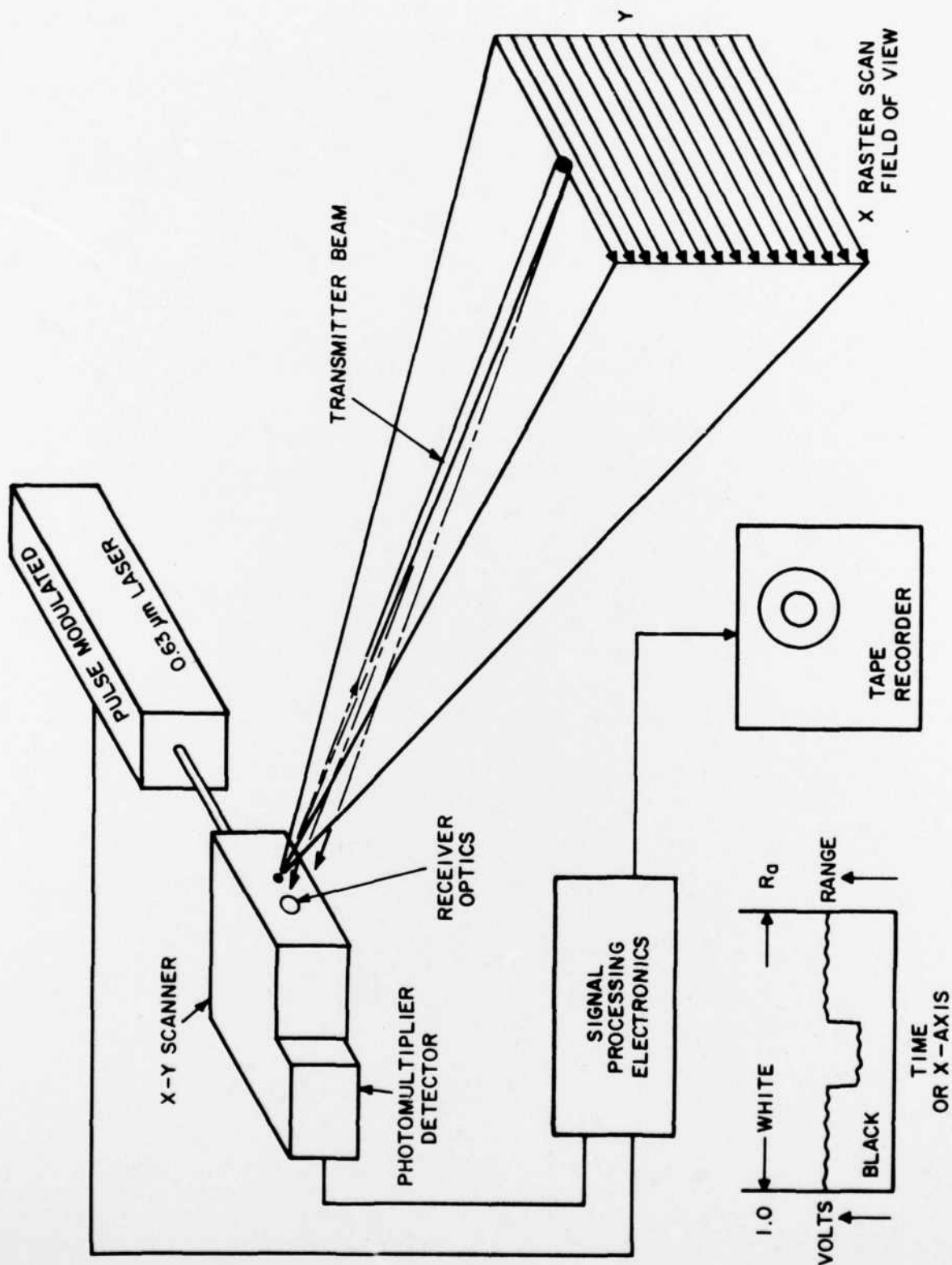- data acquisition rates in general.

GEO-CENTERS, INC.

Figure D-1. Table Top Three-Dimensional Scanner (T3DS).

GEO-CENTERS, INC.

Table-Top 3-D Scanner Specifications

SOURCE

Mode-Locked He-Ne Laser, Spectra Physics Model 125

Power:	≅50 mW maximum (~7 mW needed)
Beam Size:	2 mm
Beam Divergence:	≅1 mr
Basic Pulse Frequency:	≅79.1 MHz
Pulse Width:	1 ns
Harmonics Used:	Fund, 3 and 6
Range Ambiguity:	189.5 cm, 63.2 cm and 31.6 cm

RECEIVER

Photomultiplier ITT Type 4100, frequency response to 500 MHz

Collecting Aperture:	10 mm diameter lens
RF Electronics:	Commercial filters, mixers, amplifiers
Phase Detector:	Analog Pulse Type (from AVD-4 electronics)
RMS Range Accuracy:	~3mm for all range ambiguities

SCANNING PARAMETERS

Maximum Scanning Field of View:	14° x 14° or 480 mm x 480 mm at 2-m range
IFOV (spot size):	2 mm - 10 mm at 3-m range
Frame Rate:	1/20 seconds limited by recording
Pixels/Line:	128, 256, or 512

GEO-CENTERS, INC.

APPENDIX E
ULTRASONIC PHASE MONITORING

Using ultrasonics as a sensing technology has both positive and negative aspects. A properly configured system can be sensitive to extremely small dimensional changes, but the system generally has a very limited ability to localize or identify the source of those changes. Thus in a monitoring or inspection role, ultrasonics can be very important, but it has limited applicability in a generalized imaging role.

The following paper* discusses the use of ultrasonics in an inspection and/or monitoring role. Implied sensitivities are excellent, providing certain configuration constraints are met. In these roles, and in proximity sensing, ultrasonics will undoubtedly be a major contributor to the growth of robotic technology.

GEO-CENTERS, INC.

PHASE MONITORING FOR AUTOMATED INSPECTION, POSITIONING AND ASSEMBLY

by

SHAWN BUCKLEY, Associate Professor, Massachusetts Institute of Technology
KIM STELSON, NSF Fellow, Massachusetts Institute of Technology

ABSTRACT. Phase monitoring is a technique which can be used in several aspects of automated manufacturing: inspection, assembly and positioning. In phase monitored inspection, acoustic waves reflected off a master part are picked up by an array of microphones. The phase shift of the reflected wave in each microphone is compared to the subsequently measured values of a part to be inspected; differences in these phase signals can be interpreted as geometry differences between master and part. Similarly, the phase monitoring technique can be used for automated assembly by attaching the sound source and microphone array to a manipulator hand; the phase signals determine the relative position between the manipulator and the object being assembled. Automated positioning, on the other hand, uses similar phase signals to position objects in the proper orientation before further machining or assembly.

Phase monitoring is suitable for low production runs (where economics eliminate specialized transfer line equipment) since it can be easily recalibrated for different part shapes. PM is expected to be low cost--estimated at between \$2,000 and \$5,000 per system--because its components are all off-the-shelf speakers, microphones and microprocessors. Other advantages are high speed (several hundred measurements a second), no moving parts, and reasonable stand-off distance (several inches). Accuracy is high: changes in geometry of as small as a mil (.03mm) can be detected with present equipment.

INTRODUCTION. Inspection and assembly today are the least automated of manufacturing operations, though some automation of these operations does exist.[1] Automated inspection can be as sophisticated as Bendix's Cordax inspection machine operating under computer control;[2] automated assembly can be as sophisticated as Kawasaki's assembly of motorcycle engines. However, most shop practice is with hand-operated inspection and assembly equipment: very little automation has worked its way into these processes.

Consider automated inspection as an illustration. It has been applied most successfully to large production runs. For example in the manufacture of automobile connecting rods, precision inspection of the machined surfaces is automatic. Because of the large number of nearly identical parts, economics dictate an inspection "transfer line". The inspection machine is specially designed for each part shape. While effective for large production runs (a million parts or more) the transfer line methods are not economical for low-volume inspection. An inspection system must be flexible to inspect parts of many different shapes.

Phase monitoring is a technique which can eliminate much of the labor involved in hand inspecting and assembly for the smaller production runs where transfer line equipment is not suitable. Phase monitoring (PM) is flexible enough to be easily recalibrated for each new part shape. It is expected to be a low-cost method--estimated costs are between \$2,000 and \$5,000 per inspection station. PM techniques have other advantages of high speed (several hundred measurements a second), no moving parts and reasonable stand off distances (several inches).

Automated inspection will be used to illustrate the PM technique; assembly and positioning will be shown later to be variations of inspection. Figure 1 shows a schematic of a PM automated inspection system. An array of wave emitters and wave receivers are located around the object to be inspected: for simplicity, only a single emitter is shown. The emitter sends out acoustic waves of a constant wavelength which are reflected from the object and then picked up by the receivers. As a result of the constant wavelength emission, the output of the receivers will be a sinusoidal signal of the same frequency as the emitted wave but differing in both amplitude and phase. Inspection of the part is made possible by monitoring the phase differences between emitted and received waves.

First, a master workpiece is positioned at the inspecting station. The plurality of the phase measurements monitored at each receiver constitutes the "phase vector" of the master workpiece:

$$\theta = (\theta_1, \theta_2, \ldots \ldots \theta_n). \qquad (1)$$

A subsequent object to be inspected is positioned at the same point and orientation, and a similar phase vector:

$$\theta' = (\theta'_1, \theta'_2, \ldots \ldots \theta'_n) \qquad (2)$$

is monitored for the object. The difference in phase between the master and the object:

$$\Delta\theta = (\theta_1 - \theta'_1, \theta_2 - \theta'_2, \ldots \ldots \theta_n - \theta'_n) \qquad (3)$$

can be used to determine whether a part is within tolerance, where it isn't, and by how much. Assume for the time being that the object and master can be placed in the same orientation and position; this positioning can be either by conventional methods or by PM positioning (to be discussed later).

For many gross errors the object can be rejected if the phase difference vector is simply unrecognizable. Smaller errors require more sophistication. A sensitivity matrix, $\underline{S}$, can be determined which gives the sensitivity of each phase difference, $\Delta\theta_i$, for each dimension tolerance of the object; each element of $\underline{S}$ is the partial derivative of a particular dimension to the phase change at a particular microphone. Multiplication of the phase difference vector times the inverse of the sensitivity matrix results in a deviation vector:

$$d = \underline{S}^{-1} \Delta\theta \qquad (4)$$

The deviation vector is the amount that each dimension of the object has changed or "deviated" from that of the master workpiece:

$$d = (d_1, d_2, \ldots \ldots d_m). \qquad (5)$$

If any change is outside the tolerance for that dimension the object is rejected.
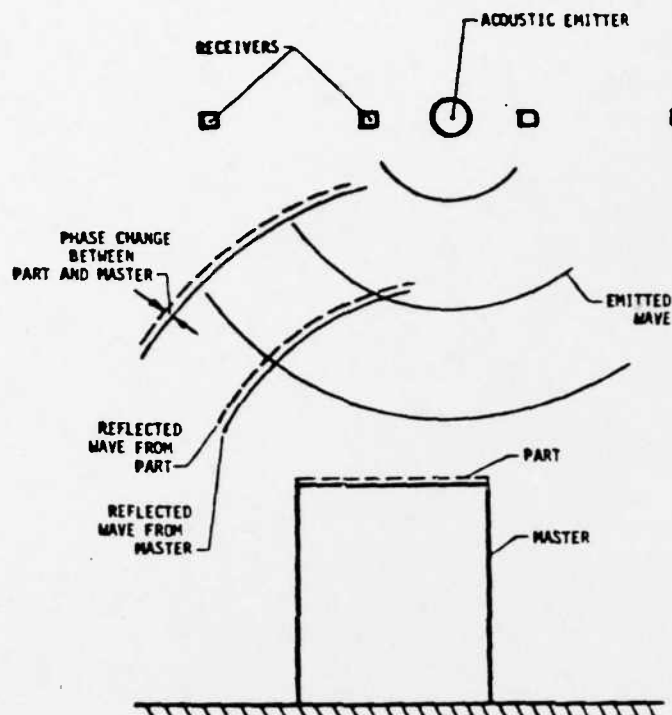
FIGURE 1   SCHEMATIC OF PM INSPECTION

To have sufficient information, the number of phases measured must be greater than or equal to the number of dimensions inspected. Redundant phase measurements are likely to provide more reliable estimates of d. Moreover, they also mitigate the possibility that $\underline{S}$ may be singular or nearly singular. In short, more phase measurements increase the probability that every dimension is well measured.

When the number of measurements exceeds the number of dimensions, $\underline{S}$ will be non-square and the inverse of $\underline{S}$ does not exist. A pseudo-inverse must be used.[3] The pseudo-inverse of $\underline{S}$ is not unique. The residual, r, is given by:

$$r = \underline{S} \, d - \Delta\theta \qquad (6)$$

The pseudo-inverse that minimized the Euclidean norm of the residual, r, is called the left minimum inverse. Since minimizing the norm of the residual is the same as minimizing the sum of the squares of the components of r, the left minimum inverse is simply a least-squares fit of the dimensions to the excess measurements. The left minimum inverse is given by:

$$\underline{S}^{LM} \equiv (\underline{S}^T \underline{S})^{-1} \, \underline{S}^T. \qquad (7)$$

Then, the deviation vector can be computed as follows:

$$d = \underline{S}^{LM} \Delta\theta \qquad (8)$$

The sensitivity matrix described above assumes linear changes in phase for small deviations. This will be true if the emitted wavelength is much greater than the largest possible deviation. A one-inch (25mm) wavelength allows deviations as large as ±.050 inch (1.3mm) to be measured while still satisfying this criterion. A one-inch (25mm) wavelength corresponds to a frequency of about 10 KHz in air-well within the realm of hi-fi equipment. However, a slightly higher frequency (e.g., 20 to 30 KHz) might be convenient to keep the emitted waves out of the audible range.

The sensitivity matrix itself is determined by a calibration procedure performed only once for each master workpiece. It involves matching the phase differences, $\Delta\theta$, with actual deviations, d, of objects with high and low tolerance limits. Once the sensitivity matrix has been determined, all subsequent objects characterized by that particular master workpiece can be inspected on-line, without the need to continually recalibrate.[4]

As you have seen, by paying special attention to phase changes of reflected acoustic waves, the deviation of a part from the master can be determined. PM assembly and PM positioning use similar methods of matrix manipulation. The simplest case is PM positioning where an object (e.g., a transmission casting) is positioned on a pallet. Automated positioning is useful in certain automated manufacturing processes when a part must be located in some known position for subsequent machining or assembly

E-3

operations. Positioning is a simple task be-
cause at most there are six possible devia-
tions corresponding to three degrees of trans-
lation and three degrees of rotation. However,
in an actual positioning operation, often only
two or three degrees are required (rotation
and XY positioning). In PM positioning, a de-
viation is defined as the error from a nominal
position rather than a change in dimension as
in PM inspection. So three degrees of posi-
tioning constitute an $\underline{S}$ matrix with only three
deviations. As long as the deviations are
relatively small compared to the wavelength
(say, 1/8 inch (3.2mm) compared to an inch
(25mm) wavelength) the same theory applies to
PM positioning as PM inspecting.

PM assembly can be thought of as many con-
tinuous PM positioning steps in succession.
Imagine an array of microphones and an acoustic
emitter attached to a manipulator hand. In
each position of the hand, the phase signals
will represent the relative position of ob-
jects close to the hand. Thus the hand can
follow some predetermined trajectory based <u>not</u>
on absolute XYZ coordinates but rather on co-
ordinates associated with the objects being
assembled. Since most current manipulators
can easily arrive within an eighth-inch (3.2mm)
of some predetermined position, PM assembly
becomes a vernier calibration movement, re-
positioning the manipulator hand relative to
the objects being assembled. For this appli-
cation phase monitoring is accurate to a mil
(.03mm), allowing accurate repositioning of
the manipulator's coordinates. Furthermore,
the PM system need not necessarily be on the
manipulator hand--it could be integrated into
the holding fixtures of the assembly system.
Since the PM technique measures <u>relative</u> po-
sitions, it makes no difference whether it is
on the fixture "seeing" the hand.

Phase monitoring, then, is a way of find-
ing "deviations" in objects. The deviations
can either be geometry changes from object to
object (inspection), changes in a single ob-
ject's position (positioning) or changes in
the relative positions between objects (assem-
bly). The use of phase as a measurement vari-
able has inherent advantages. By analogy,
consider a tape measure which measures distance
giving only <u>fractions</u> of an inch, not the <u>num-
ber</u> of inches. However, this tape measure <u>can</u>
measure these fractions very accurately, say,
to one part in a thousand. Suppose you were
to use this tape measure to find the length
of an object whose nominal measurement was 10
inches (.4m). If you measured .001 inches
(.03mm) you would know it's relative tolerance
to one part in ten thousand, even though your
measurement was only accurate to one part in a
thousand. Similarly, phase can be measured to
one part in a thousand even though the distance
between object and the microphone array may be
several wavelength. In most practical circum-
stances measuring the proper phase but the
wrong wavelength (9.001 instead of 10.001 in
the tape measure analogy) won't occur if the
phase changes are small compared to a wave-
length. In inspection, positioning and assem-
bly, all changes are small giving advantages
to the inherent sensitivity of phase measure-
ment.

EXPERIMENTAL VERIFICATION. An experimental
apparatus was built to verify the feasibility
of phase monitoring. A simple, two dimension-
al PM positioning system was investigated. To
insure general results, a complex part was used
for the study: an automobile carburetor. The
carburetor was mounted on "translation stages"

to allow accurate movement in two orthogonal
directions. Micrometer heads on the transla-
tion stages allowed the carburetor to be posi-
tioned accurately.

Moving the carburetor slightly in each of
two directions ($\Delta x, \Delta y$) caused phase changes in
each of four microphones ($\Delta\theta_1, \Delta\theta_2, \Delta\theta_3, \Delta\theta_4$).
The resulting sensitivity matrix was:

$$\underline{S} = \begin{bmatrix} \dfrac{\Delta\theta_1}{\Delta x} & \dfrac{\Delta\theta_1}{\Delta y} \\[2mm] \dfrac{\Delta\theta_2}{\Delta x} & \dfrac{\Delta\theta_2}{\Delta y} \\[2mm] \dfrac{\Delta\theta_3}{\Delta x} & \dfrac{\Delta\theta_3}{\Delta y} \\[2mm] \dfrac{\Delta\theta_4}{\Delta x} & \dfrac{\Delta\theta_4}{\Delta y} \end{bmatrix} = \begin{bmatrix} -0.645 & -0.555 \\[2mm] -0.590 & 1.045 \\[2mm] -0.325 & 0.175 \\[2mm] 0.085 & 1.380 \end{bmatrix} \dfrac{degrees}{mil} \quad (9)$$

General   Experimental
Expression   Data

Each element of the $\underline{S}$ matrix is, in essence,
the partial derivative of each microphone's
phase difference with respect to each deviation.
Given the sensitivity matrix, any small posi-
tion change of the carburetor in the x or y di-
rection could be found by monitoring the phase
differences at each microphone.

To illustrate the matrix technique, con-
sider the case where the carburetor was moved
10 mils (.25mm) in <u>both</u> the x and y directions.
The measured phase change vector for this com-
bined movement was:

$$\Delta\theta = \begin{bmatrix} -18.0 \\ 9.1 \\ -3.3 \\ 9.9 \end{bmatrix} \quad degrees \quad (10)$$

To examine the improvement in accuracy
achieved by using redundant phase measurements,
comparisons were made among three different
cases: 1) using only the first two microphones,
2) using only the last two microphones and 3)
using all four microphones. The first two
cases involve the inversion of a square sensi-
tivity matrix; the last case involves pseudo-
inversion of a non-square sensitivity matrix.

Using only the first two microphones (i.e.,
the top four elements of Equation 9 and the top
two elements in Equation 10):

$$\underline{S} = \begin{bmatrix} -0.645 & -0.555 \\ -0.590 & 1.045 \end{bmatrix} \dfrac{degrees}{mil} \quad (11)$$

$$\begin{bmatrix} d \\ x \\ d \\ y \end{bmatrix} = \underline{S}^{-1}\Delta\theta = \begin{bmatrix} -1.045 & -0.555 \\ -0.590 & 0.645 \end{bmatrix} \cdot \begin{bmatrix} -18.0 \\ 9.0 \end{bmatrix} = \begin{bmatrix} 13.8 \\ 16.5 \end{bmatrix} mils \quad (12)$$

Since the actual movement of the carburetor was:

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 10.0 \\ 10.0 \end{bmatrix} \quad \text{mils} \qquad (13)$$

using only the first two microphones did not provide a very good estimate of the carburetor movement.

Similarly, using only the last two microphones (the bottom four elements of Equation 9 and the bottom two elements in Equation 10), gave a poor estimate of the actual movement.

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \begin{bmatrix} 6.1 \\ 6.8 \end{bmatrix} \quad \text{mils} \qquad (14)$$

However, using all four microphones (all the elements of Equations 9 and 10) gives:

$$\underline{S}^{LM} = (\underline{S}^T\underline{S}^{-1})\,\underline{S}^T = \begin{bmatrix} -0.750 & -0.644 & -0.377 & 0.137 \\ -0.186 & -0.297 & -0.062 & 0.418 \end{bmatrix} \frac{\text{degrees}}{\text{mil}} \qquad (15)$$

$$\begin{bmatrix} d_x \\ d_y \end{bmatrix} = \underline{S}^{LM}\Delta\theta = \begin{bmatrix} 10.2 \\ 10.4 \end{bmatrix} \quad \text{mils} \qquad (16)$$

For this case, pseudo-inversion provides a substantial improvement in the estimate of the deviation from the actual value of 10 mils (.25mm) in each direction. This will not be universally true: the pseudo-inverse provides only a most likely estimate of the actual deviations. If the point picked for calculation were near a relative maxima or minima in the phase variation, higher order effects might be more influential, causing decreased accuracy. This possibility was avoided for the purpose of the example shown.

In PM inspection and PM assembly, the same matrix techniques are valid. As mentioned before, in PM inspection the phase signals of a part are compared to a master. In PM assembly the phase signals correspond to relative positions between the objects being assembled. However, the distinction between the various PM techniques is blurred. For example, PM inspection requires that subsequent parts be placed in exactly the same position so that phase differences correspond only to differences between part and master. However, it's possible to optimally position the part using PM positioning and then subsequently PM inspect it. In fact, if there are only small errors in its position, the part can be inspected without repositioning: the matrix technique can "wash out" the effect of misalignment.

RESOLUTION. While the previous discussion has shown how phase monitoring using several microphones can measure geometry changes, the accuracy associated with the measurement has not been given. How small a defect can be detected?. Can microphones be placed in optimal positions to detect tolerances best? These and similar questions can be answered by turning to acoustic theory.

In general, geometry changes or deviations can be divided into two sizes: 1) changes in objects whose surface area is greater than a wavelength squared and 2) changes in objects whose surface area is smaller than a wavelength squared. Recall the PM systems use a wavelength about an inch (25mm) long, so a wavelength squared has about one square inch ($6 \times 10^{-4}m^2$) surface area. For big surfaces, the ray-tracing method has been found to give the most insight into applications of phase monitoring. PM positioning and PM assembly usually involve these large surfaces. In assembly, the "surfaces" are the objects and fixtures "seen" by the microphone array in the manipulator hand. If the array were mounted in the holding fixture, the "surfaces" would be those of the manipulator hand itself as it approached the objects being assembled. In PM positioning the "surfaces" are just those of the object being positioned. In both PM assembly and PM positioning small relative movement of objects with large surface area is the important theme.

PM inspection, on the other hand, usually deals with geometry changes of small surfaces. Huygen's diffraction theory must be used for analysis since the ray-tracing method is not appropriate. For example, detecting the presence or absence of a machine screw on an assembled part must use diffraction theory because the surface area of the machine screw is so small.

In analyzing large surfaces, the ray-tracing method assumes three types of acoustic "rays" are combined in each microphone: the direct ray, reflected rays and diffracted rays. The direct ray is the pressure signal coming directly from the wave emitter to a particular microphone. Reflected rays reach the microphone by being reflected from the object's surface; geometric reflection is assumed. Diffracted rays are emitted from each sharp corner. Figure 2 shows direct, reflected and diffracted rays converging on a microphone in a very simple geometry: a two dimensional wedge. These waves are actually time-varying pressure waves of different amplitude and phase shift which can be combined vectrially using phasor diagrams (see Figure 2).

To investigate the validity of the ray-tracing technique, one of the authors (Stelson) has studied ray-tracing experimentally. He used a two-dimensional wedge geometry as shown in Figure 2. An omni-directional source and microphone were held fixed while the wedge was moved in two dimensions, simulating PM positioning of a simple object. The results of the experiment are shown in Figure 3. The vertical scale shows absolute phase shift as received by the microphone while the horizontal axis shows the wedge's position. A theoretical prediction of the phase at the microphone is shown along with experimental data. The prediction used a computer program that computes the combined effects of direct, reflected and diffracted rays. The theory agrees well with the experiment, verifying ray-tracing for determining absolute phase. The errors between theory and experiment in Figure 3 are not indicative of the precision with which PM can measure changes in geometry; they represent the differences between an ideal wedge and the real wedge. Note that the data presented is merely representative; other tests showed similar correlation with theory.[5]
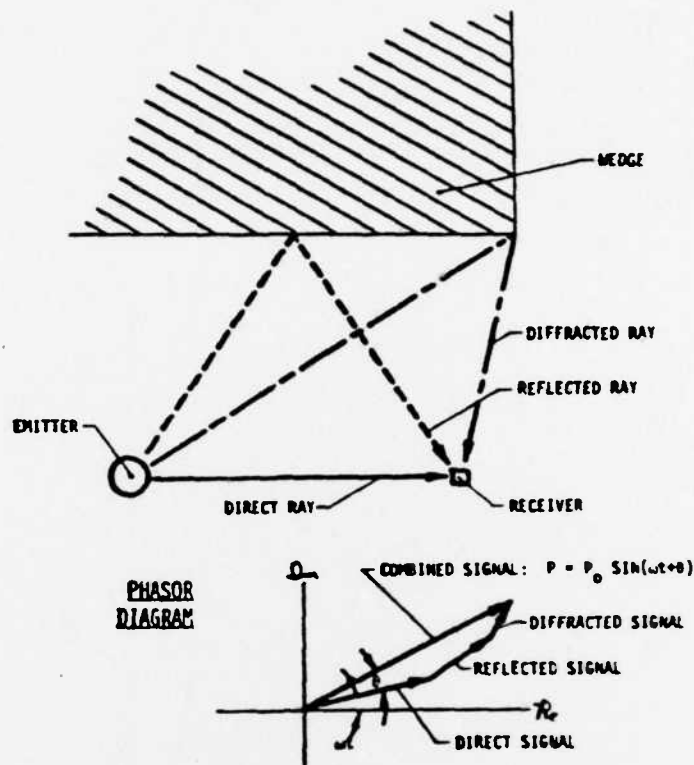
E-5
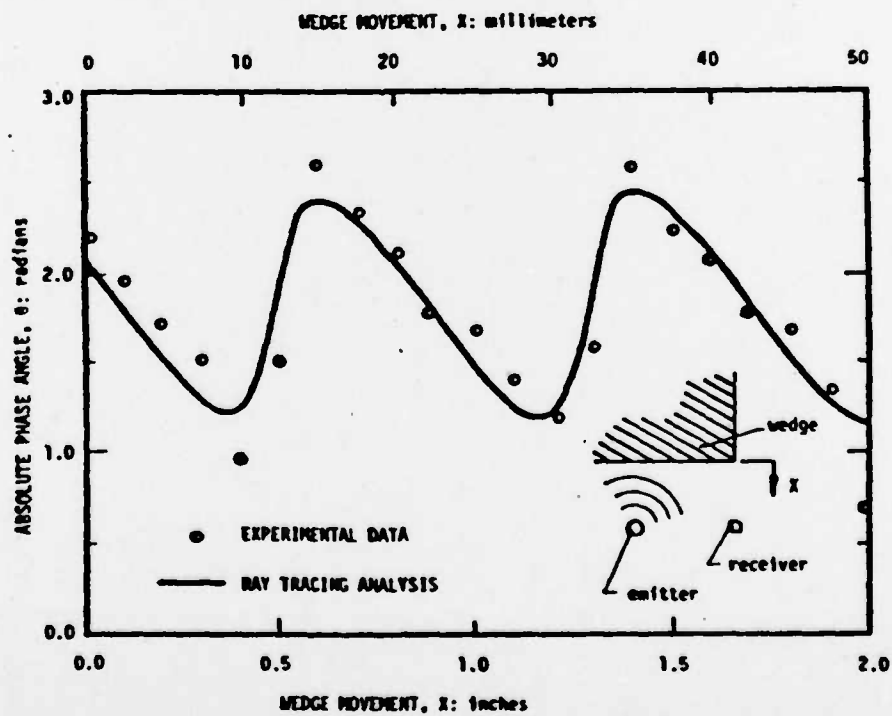
FIGURE 2   RAY TRACING FOR A WEDGE

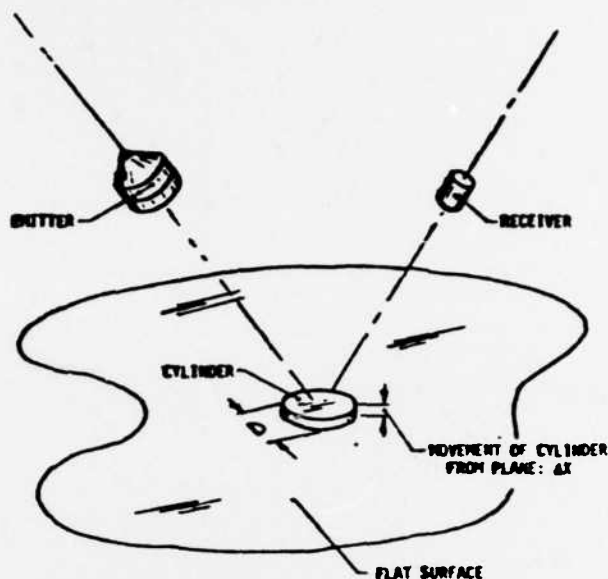FIGURE 3   RAY TRACING: ANALYSIS AND EXPERIMENT FOR A WEDGE

E-6

FIGURE 4   SCHEMATIC OF DIFFRACTION EXPERIMENT

For objects with small surface area, a variation of Huygen'a diffraction theory ia appropriate. In this theory, pressure waves atriking an area-increment much smaller than the wavelength aqusred are assumed to disperse omni-directionally awsy from the plane of the surface. The pressure aignal from the aurface becomes the phasor combination of aignals from each area-increment. Each area-increment behaves as if it were a small pressure aource located on the object's aurface. The phasor sum of these sources is combined with the direct pressure aignsl coming from the microphone, as well as with reflected aignals coming from other nearby objects.

The diffrsction theory vss experimentally verified by the senior author. Cylinders were pushed through close-fitting holes in a flat surface to aimulate the effect of a small defect in the aurface. The cylinders began flush with the aurface and were raised an eighth inch (3.2mm) above the surface as shown in Figure 4. The phase of the microphone aignal vas primarily determined by direct vsvas from the emitter and to a lesser degree by reflected wsves from the flst surface. This dominant signal was changed slightly by the contribution of the signal diffracted from the "defect".

Results of the small-surfsce experiment are shown in Figure 5. The vertical axis ahows the phase change (the difference in the microphone signal before and after the cylinder movement) the horizontsl acale shows the diameter of the cylinder. Also shown in Figure 5 are the results of a computer program using the Huygen's diffraction principle. Since the exact phase chsnge depends on the ration the direct and reflected components compared to the diffracted component (which is very difficult to determine analytically), the theoretical results have been normalised to the experimental results for the half-inch (12.7mm) diameter cylinder.

Diffraction theory predicts the phase change due to displscement fcr each tiny srea-increment of the cylinder's face will give approximately the aame phase aignal change at the microphone. Since a half-inch (12.7mm) diameter cylinder bas four times as many area-increments as a quarter-inch (6.4mm) diameter cylinder, one would expect its phase signal to change about four times as much. The experiment verifies this prediction. In general, the phase change is proportionsl to the area of a small surface being inspected.

What do these experiments mean in terms of a real inspection, positioning or assembly tssk? Our present equipment can detect phase changes of about .001 cycles (.006 rad). This means for big surfaces (ususlly a positioning or assembly task), an object can be located to within a mil (.03mm) using a one-inch (25mm) wavelength. For small aurfsces (ususlly inspection tasks), the concept of a "defect volume" is appropriate; it is simply the volume of a small defect in a psrt being inspected. Our present equipment can detect defect volumes as small as .002 cubic inches (30mm^3). Thus s half-inch (12.7mm) dismeter defect only 10 mils (.25mm) high csn be detected, but an eighth-inch(3.2mm) diameter defect must be an eighth of an inch (3.2mm) high to be detected. Diffraction sees bumps and holes as thc ssme so the resolution volume applies to "positive" and "negstive" defects. Note that these resolutions--1 mil (.03mm) positioning for lsrge surfsces and .002 in^3 (30mm^3) defect volumes for small surfaces--only applies to our preaent inch-long (25mm) wsvelength. By using higher frequency sound greater resolution is possible, but at the expense of amaller atandoff distance.

COMPARISON WITH OTHER SYSTEMS. Two types of syatems are currently being used for automated inspection, positioning and assembly. They can loosely be referred to as "point-to-point" techniques and "video" techniques.
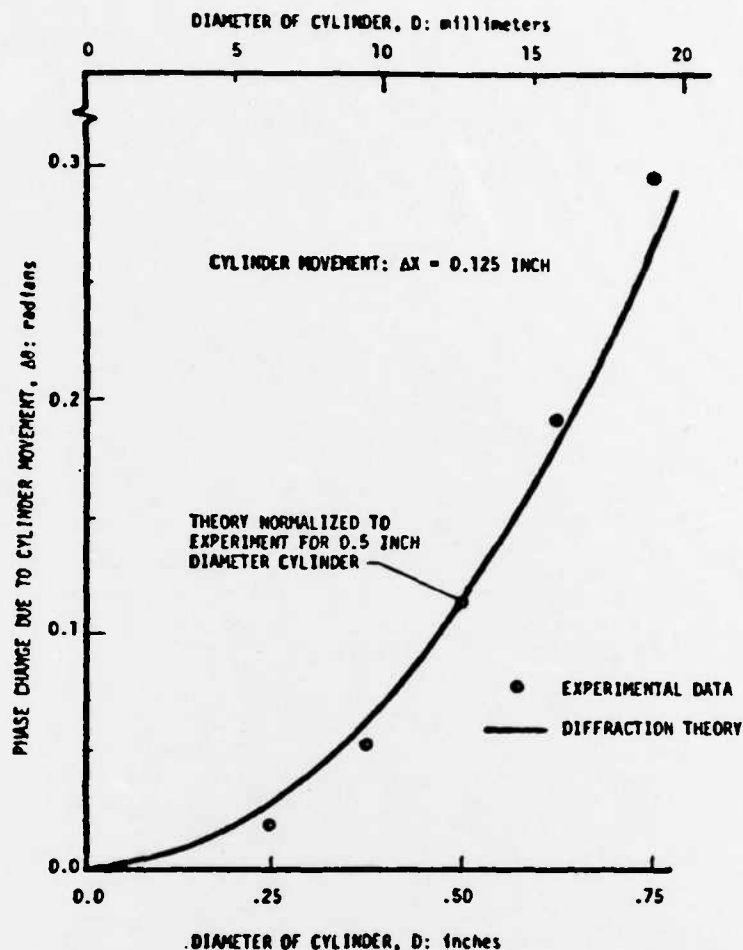
E-7

FIGURE 5   RESULTS OF DIFFRACTION EXPERIMENT

Point-to-point machines, used primarily for inspection, are laden with mechanical or electronic sensors which probe the part for any anomalies. The sensors are mechanically moved into close proximity of the surface being inspected, since the sensors require small standoff distances. Motion is controlled by servo-motors or stepper motors similar to the motion in an N/C machine tools; however, a sensor replaces the cutter. As their name implies, point-to-point machines make measurements at individual points one after the other.

On the other hand video techniques generally use either linear or area arrays of light detectors. An ordinary television camera would fall into this category (although most recent devices use solid-state detectors). Coupled with a microprocessor, video techniques have automated both inspection and assembly processes.

Compared to point-to-point machines, PM is both faster and has averaging capability. Since the point-to-point machines measure sequentially taking perhaps a second for each point, a part with many measurements can take a minute or more for inspection. A PM system makes all its measurements in parallel, rather than in series, and so is inherently faster.

A PM inspection machine could measure 10 or 20 dimensions in less than a second. A point-to-point machine is only sensitive to the actual point being measured (it has no way of knowing if a defect exists at other than the measured point). The capability of averaging many measurements over an entire surface is often necessary when working with objects such as sand castings. Although the surface of a casting has a rough surface with many flaws, the PM technique can detect the nominal location of the surface.

Video techniques are just as fast as PM, but they seem to be limited to two-dimensional shapes. The variable surface reflectance of three-dimensional objects often gives spurious results to video techniques; when an object can be presented as a silhouette, video is at its finest. PM systems inherently use three-dimensional information of objects. Furthermore, in transforming the acoustic waves into phase signals, the shape of the object becomes unimportant: turbine blades are no more difficult to measure than spheres or cubes. Although both PM and video use microprocessors, PM has the advantage of much smaller microprocessor memory size and less computation time. Whereas video techniques store and manipulate information associated with a picture of the entire object, in PM only the barest minimum of information is

stored or manipulated. For example, using 10 microphones means only ten 10-bit words need be stored and manipulated.

Economics are another way to compare the various systems. A point-to-point inspection machine is usually in the $50,000 class: while both video and PM systems are estimated to cost about $5,000 in production. In terms of cost, one would expect little difference between a video system and a PM system: both use a microprocessor computer with auxiliary source and detectors. In video, a light source and optical detectors are used; in PM, a sound source and microphones are used. As the price of microprocessors is reduced, so will be the cost of PM and video systems: $1,000 per unit prices are feasible.

SUMMARY. Phase monitoring can be used for automated inspection, positioning and assembly processes. Without moving parts and with relatively large standoff distances, PM can detect slight differences between a master and another part, between an object and its nominal position and between two objects being assembled. The system is flexible, permitting easy recalibration for objects of different shape. Estimated costs for the system are expected to be in the $2,000 to $5,000 range. Present digital equipment can make several hundred measurements a second.

Acoustic theory has been successfully applied to the PM process, resulting in a high degree of understanding of how geometric changes are transformed into phase changes. Accuracies in detecting changes on two classes of surfaces have been both analytically and experimentally investigated, with good agreement in the results. For large surfaces--areas more than a wavelength squared--ray-tracing algorithms are appropriate. For small surfaces--areas less than a wavelength squared--diffraction theory gives insight into the PM process. With present equipment, changes in position of large surfaces as small as 1 mil (.03mm) can be detected. The same equipment can detect a defect in an object as small as an eighth-inch (3.2mm) on a side.

NOMENCLATURE.

d   Deviation or geometric change

d_i Deviation of the $i\underline{th}$ dimension

D   Diameter of cylinder

p   Pressure

r   Residual error in measurement

$\underline{S}$ Sensitivity matrix

$\underline{S}^{-1}$ Inverse of sensitivity matrix

$\underline{S}^{T}$ Transpose of sensitivity matrix

$\underline{S}^{LM}$ Left minimum inverse of sensitivity matrix

t   Time

x   Coordinate direction

y   Coordinate direction

Φ Vector of phases received by all microphones

Φ_i Phase of the $i\underline{th}$ microphone

$\Delta\Phi$ Vector of phase differences received by all microphones

ω Frequency of acoustic waves

Δx Small movement in x direction

Δy Small movement in y direction

REFERENCES

1. Harrington, J., A. D. Little Co., presented to Materials Engineering Congress, American Society of Metals, 1971 (reprinted in Automation, Vol. 20., No., January 1973).

2. Schaffer, G., "Better Molds Via Computer," American Machinist, Vol. 114, No. 15, July 27, 1970.

3. Takahashi, Y. Rabins, M. and Auslander, D., "Control," Addison-Wesley, Menlo Park, Ca., 1970.

4. Buckley, S., "Phase Monitored Inspection," Proceedings of the Third North American Metal Working Research Conference, Pittsburgh, Pa., May 1975.

5. Stelson, K., "Phase Monitored Inspection: Theory and Application," M.S. Thesis, Mechanical Engineering Dept., M.I.T., May 1977.